

```

Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGGTGAACCTGACCTCTCTACATGCCCCATCTGGAAGTTTCCAGATGAGAG 1860
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 1861 GGCATGCGCAGCCTTGCCCATCAACTGCACCCACCTCTGTGTGGACCTGGATGACAAG 1920
Qy 641 GlyCysProAlaGluGlnAtqAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCCGAGCAGAGAGCCCTCTGACGTCCTATCATCTCTGCGGTGGTGGC 1980
Qy 653 ----- 653
Db 1981 ATTCTGTGTCGTGGTCTTGGGGTGTCTTTGGGATCCTTCATCAAGCGACGCCAGCAG 2040
Qy 653 ----- 653
Db 2041 AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAACGGAGCTGGTGGAGCCGCTG 2100
Qy 653 ----- 653
Db 2101 ACACCTAGCGAGCGATGCCAACCCAGCGCAGATCGGATCTCTGAAAGACAGCGAGCTG 2160
Qy 653 ----- 653
Db 2161 AGGAAGGTGAAGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAAGGCACTGGATC 2220
Qy 653 ----- 653
Db 2221 CCTGATGGGAGAATGTGAATTTCCAGTGGCCATCAAAAGTGTGAGGGAAACACATCC 2280
Qy 653 ----- 653
Db 2281 CCCAAACCAACAAAGAAATCTTAGACGAAGCATACGTGATGCGTGTGGGCTCCCA 2340
Qy 653 ----- 653
Db 2341 TATGTCTCCGCCCTTCTGGCATCTGCTGACATCCACGCTGACAGCTT 2400
Qy 653 ----- 653
Db 2401 ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCCGGACGCCCTGGGCTCCAG 2460
Qy 653 ----- 653
Db 2461 GACCTGCTGAACGTGTATCCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTCGG 2520
Qy 653 ----- 653
Db 2521 CTGTTACACAGGACTTGGCGGCTCGGAAGCTGCTGCAAGAGTCCCAACCATGTCAA 2580
Qy 653 ----- 653
Db 2581 ATTACAGACTTGGGCTGGCTCGGCTGCTGGACATTCAGCATTCAGCAGACAGATACCATGCAGAT 2640
Qy 653 ----- 653
Db 2641 GGGGCAAGGTGCCATCAAGTGGATGGCGCTGGAGTCCATTTCTCCCGCGGGTTCCAC 2700
Qy 653 ----- 653
Db 2701 CACCAGATGATGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC 2760
Qy 653 ----- 653
Db 2761 AAACCTTACGATGGATCCAGCCCGGGAGATCCCTGACCTGCTGGAAAGGGGAGCGG 2820
Qy 653 ----- 653
Db 2821 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTGTTCAATGTGGATG 2880

```

```

653 ----- 653
2881 ATTCACTCTGAATGTCGGCCAAGATTCCGGAGTTGGTGTCTGAATTTCTCCCGCATGGCC 2940
654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
2941 AGGACCCCCAGCCTTTGTGTCATCAGAAATGAGGACTTTGGGCCAGCCAGTCCCTTG 3000
665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
3001 GACAGACCTTCTTACCCTCTACTGCTGAGGACCATGACATGGGGGAGCTGGTGGATGCT 3060
685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
3061 GAGGAGTATCTGGTACCCAGCAGCGCTTCTTCTGTCACAGACCTGCCCGGCGCTGG 3120
705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
3121 GGCATGTCACCAACAGCAGCCGACGCTCATCTACCAGGAGTGGCGGTGGGACCTGCACA 3180
725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
3181 CTAGGGCTGGAGCCCTCTGAAGAGAGGCCCCCAGGTCTCCACTGGCACCCCTCCGAGGG 3240
745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
3241 GCTGGCTCCGATGTATTGATGGTGACCTGGGAATGGGGCAGCCAAAGGGCTGCAAGC 3300
765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
3301 CTCCCCACATGACCCCGCCCTCTACAGCGGTACAGTGGAGGACCCACACAGTACCCCTG 3360
785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
3361 CCCTCTGAGACTGATGCTAGCTTGGCCCCCTGACCTGCAGCCCCCAGCTGGAATATGTG 3420
805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
3421 AACCCAGCAGATGTCGGCCCCAGCCCTTCGCCCGAGAGGGCCCTCTGCCCTGCTGCC 3480
825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
3481 CGACCTGCTGGTGCACTCTGGAAGGCCCAAGACTCTCTCCCGAGGGAAGAAATGGGGTC 3540
845 ValLysAspValPheAlaPheGlyGlyValAlaGluAsnProGluTyrLeuThrProGln 864
3541 GTCAAGACCTTTTGGCTTTGGGGTGCCGTGGAGAACCCCGAGTACTTGACACCCAG 3600
865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
3601 GGAGGAGCTGCCCTCAGCCCCACCTCTCTGCTTCCAGCCAGCCCTTCGACAAACCTC 3660
885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
3661 TATTATGGGACCAAGGACCCAGAGCGGGGGCTCCACCCAGCAGCTTCAAGAGGGACA 3720
905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
3721 CCTACGCACAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG 3765

```

# RESULT 2

US-09-930-125-1

```

: Sequence 1, Application US/09930125
: Publication No. US2002019329A1
: GENERAL INFORMATION:
: APPLICANT: Hand-Zimmerman, Susan
: APPLICANT: Cheever, Martin A.
: APPLICANT: Foy, Teresa M.
: APPLICANT: Lodes, Michael J.
: APPLICANT: Kalos, Michael D.
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Vedvick, Thomas S.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS

```

```

: TITLE OF INVENTION:  OF HER-2/NEU-ASSOCIATED MALIGNANCIES
: FILE REFERENCE:  210121.544
: CURRENT APPLICATION NUMBER:  US/09/930,125
: CURRENT FILING DATE:  2001-08-14
: NUMBER OF SEQ ID NOS:  25
: SOFTWARE:  FastSeq for Windows Version 3.0
: SEQ ID NO: 1

```

Alignment Scores:

Pred. No.:	0	Length:	3768
Score:	4900.00	Matches:	919
Percent Similarity:	73.23%	Conservative:	0
Best Local Similarity:	73.23%	Mismatches:	0
Query Match:	96.49%	Indels:	336
DB:	10	Gaps:	1

SEQ6 (1-919) x US-09-930-125-1 (1-3768)

1	MetGluLeuAlaAlaLeu	CysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla	20
1	ATGAGACTGGGGCTTCTGCGCTGGGGCTCTCTCTGCTTGGCCCTTGGCCCGGAGCG	60	
21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40	
61	GGAGCACCCACAGTGTGCACCGGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG	120	
41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlyAsnLeu	60	
121	ACCACCTGGACATCTCCGCACCTCTACCAGGCTGCCAGGTGGTGACGAGGAAACCTG	180	
61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal	80	
181	GAATCACCCTACCTGCCACCACTGCCAGCTGTCTCTCTGAGGATATCCAGGAGGTG	240	
81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100	
241	CAGGGCTACGTCTCATCGCTCACAACTGAGGCAGGTCCCACTGCAGAGGCTCGCG	300	
101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120	
301	AITGTGCGAGGCACCACTCTTGAGGACAACATATCCCTGCGCTGTGTAGACAATGA	360	
121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140	
361	GACCCCTGAAACATAACACCCCTGTACAGGGGCTCCCCAGGAGGCTCGCGGAGCTG	420	
141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160	
421	CAGCTTCGAAGCCTCACAGAGATCTGAAAGAGGGGTCTGTATCCAGCGGAACCCCCAG	480	
161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla	180	
481	CTCTGCTACCAAGCACAGATTTTGGAGGACATCTCCACAAGAACACACAGCTGGCT	540	
181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200	
541	CTCACACTGATAGACACCAACCGCTCTCGGGGCTGCCACCCCTGTCTCCGATGTCTAAG	600	
201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220	
601	GGCTCCCGCTCTGGGGACAGAGTCTTGAGGATTTGTCCAGAGCCTGACGGCACTGTCTGT	660	
221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240	
661	GGCGGTGGCTGTGCCGCTGCAAGGGGCACCTGCCACTCACTGTCTGCCAAGAGCAGTCT	720	





```
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-313-644-1
```

## Alignment Scores:

```
Pred. No.: 0 Length: 3768
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 12 Gaps: 1
```

SEQ6 (1-919) x US-10-313-644-1 (1-3768)

```
QY 1 MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
DB 1 ATGGAGCTGGCGCTTGGCGCTGGGGCTCCTCGCCCTCTTGCCTCCCGAGGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
DB 61 GCGAGCACCCAAAGTGTGACCGGCACAGACATGAAGCTGCGGCTCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
DB 121 ACCCAGCTGGACATGCTCGCCACCTCTACAGAGGCTGCCAGTGTGTCAGGGAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
DB 181 GAACTCACCTACCTGCCACCAATGCCAGCTGTCTCTTCTCCAGGATATCCAGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
DB 241 CAGGGCTAGTGCTCATGCTCATCAACCAAGTGAAGGAGGTCCTCCACTGCAGAGGTGGG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
DB 301 ATTGTGCGAGGACCCAGCTCTTTGAGGACAATATGCCCTGCGCGTGTAGACAAATGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
DB 361 GACCCGCTGAACAATACCACTCCCTGTCAAGGGGCTCCCAAGAGGCTCGCGGAGGTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValValLeuIleGlnArgAsnProGln 160
DB 421 CAGCTTCGAGGCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCGAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
DB 481 CTCCTGCTACCGACAGCATTTGTGGAAGGACATCTCCACAAGAACCAACAGTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
DB 541 CTCACACTGATAGACCAACACGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAA 600
QY 201 GlySerArgCysTrpGlyGluSerGluAspCysGlnSerLeuThrArgThrValCys 220
DB 601 GGCTCCCGTGTCTGGGAGAGAGTCTGAGGATTGTACAGAGCTCAGCGCACTGTCTGT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCys 240
DB 661 GCCGCTGGCTGTCGCCGTGCAAGGGCCACTGCCCACTGACTGTGCTCAGCAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
DB 721 GCTGCCGCTGACCGGGCCCCAAGCACTGTGACTGCTGCGCTGCCCTCCACTTCAACCA 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
DB 936
```

```
DB 781 AGTGGCATCTGTGAGCTGCACCTGCCAGCGCCCTGGTCACTACAACACAGACAGTGTGAG 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
DB 841 TCCATGCCCAATCCCGAGGGCGGTATACATTCGGGGCCAGCTGTGTACTGCTGCTGCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
DB 901 TACAACCTACCTTCTACGGACGTGGATCTCGACCTCGCTGCGCCCTGCACAACCAA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
DB 961 GAGGTGCACAGCAGAGATGGAACACACAGCGGTGTGAGAGTGCAGAAGCCCTGTGCCGA 1020
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
DB 1021 GTGTCTATGTCTGGGCATGGAGCACTTGCAGAGGTGAGGGCAGTTACCACTGCCAAT 1080
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
DB 1081 ATCCAGGAGTGTGCTGGCTGCAAGAAGATCTTTTGGAGCTTGGCATTTCTGCCGGAGAC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
DB 1141 TTTGATGGGGACCCAGCTCCCAACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGTGT 1200
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
DB 1201 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGGCCGAGACGCTGCCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
DB 1261 GACCTCAGCGCTTCTCAGAACCTGCAAGTAATCCGGGGACGAATCTGCACAATGGCGC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
DB 1321 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTCGCTCACTAGGGAA 1380
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
DB 1381 CTGGCAGTGGACTGSCCTCATCCACCATAACACCCACCTCTGCTTCTGTCACACGGTG 1440
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
DB 1441 CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTGTCTCCACTGCCAACCGGCCA 1500
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
DB 1501 GAGGACGAGTGTGTGGCGAGGGCTTGGCTGCCACCACTGTGCGCCGAGGGCACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnLucys 540
DB 1561 TGGGCTCCAGGGCCCAACCCAGTGTGTCAACTGCAGCCAGTTCTTTCGGGGCCAGAGTGC 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
DB 1621 GTGGAGGAATCCGAGTACTGCAAGGGCTTCCCGAGGGAGTATGTGAATGCCAGGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
DB 1681 TTGGCGTCCACCCCTGAGTGTGAGCCCAAGAAATGGCTCAGTACCTGTGTGGACCGGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
DB 1741 GCTGACCACTGTGTGGCTGTGCCCACTATAAGGACCTCCCTTCTGCTGCGCCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
DB 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
DB 1861 GCGCATGCCAGCTTGGCCCACTCACTGCCCACTCTCTGTGTGACCTGTGATGACAAG 1920
```



## Alignment Scores:

Pred. No.:	0	Length:	4473
Score:	4900.00	Matches:	919
Percent Similarity:	73.23%	Conservative:	0
Best Local Similarity:	73.23%	Mismatches:	0
Query Match:	96.49%	Indels:	336
DB:	11	Gaps:	1

SEQ6 (1-919) x US-09-441-411-5 (1-4473)

QY	1	MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla	20
DB	175	ATGGAGCTGGCGGCTTGGCGCTGGGGCTCCCTCGCCCTTGGCCCCCGGAGCC	234
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
DB	235	CGAGACCAAGTGTGACCGGCACAGACATGAAGCTGCGCTCCCTGCCAGTCCCGAG	294
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60
DB	295	ACCCACCTGGACATGCTCCGCCACCTCTACCAAGGCTGCCAGGTGGTCAGGGAACCTG	354
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
DB	355	GAACCTACCTACTGCCCCACCAATGCCAGCCTGTCTTCTCGAGGATATCCAGAGGTG	414
QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
DB	415	CAGGGCTACGTCTCATCGCTCACAAACCAAGTGAGGAGTCCCACTGACAGAGGTGGG	474
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
DB	475	ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGCGCTGTAGCAATGGA	534
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
DB	535	GAACCGCTGAACATATACACCCCTGTACAGGGGCTCCCCAGAGGCTCGGGAGCTG	594
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160
DB	595	CAGCTCGAAGCTCACAGAGATCTTGAAGAGGGGTCTTGATCAGCGGAACCCCCAG	654
QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla	180
DB	655	CTCTGCTACGAGACAGATTTGTGGAAGGACATCTTCCACAAGAACAACAGGTGGCT	714
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
DB	715	CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTA	774
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
DB	775	GGCTCCCGCTGTGGGAGAGATCTGAGGATTTGTACAGGCTGCAGAGCTGACGCGCA	834
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCys	240
DB	835	GCGGTGGCTGTGCGCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGT	894
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
DB	895	GCTGCGGCTGACCGGCCCCAAGCACTCTGACTGCTGGCTGCCCTCCCTCAACCCAC	954
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
DB	955	AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTACACACACACACAGTTGAG	1014
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
DB	1015	TCCATGCCCAATCCCGAGGCGGGTATACATTGCGCGCCAGCTGTGTGACTGCCGTGCC	1074
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
DB	1075	TACACTACTTTTACGAGACGTGGATCTTCACCCCTGTCTGCCCCCTGCACACCA	1134

QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
DB	1135	GAGGTGACAGCAGAGATGTAACACACAGCGGTGTGAGAAGTGACGAAGCCCTGTGCCCGA	1194
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
DB	1195	GTGTGCTATGTTGTGGCATGGAGCACTTGCAGAGGTGAGGCAGTTACCAAGTGC	1254
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
DB	1255	ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC	1314
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGlnLeuGlnValPhe	400
DB	1315	TTTGTATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGACAGTCCAAGTGT	1374
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
DB	1375	GAGACTCTGGAAGAGATCACAGTTACCTATATCATCTCAGCATGGCCGACAGCCTGC	1434
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
DB	1435	GAACCTACGCTCTCCAGAACCTGCAAGTAAATCCGGGGACGAATCTTGCCCAATGGCGC	1494
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
DB	1495	TACTCGTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTCGCTCACTAGGGAA	1554
QY	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
DB	1555	CTGGCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTGTCGACACGGTG	1614
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
DB	1615	CCCTGGGACCACTCTTTCGGAAACCCGACCAAGCTCTGCTCCACTGCCAACCGGCCA	1674
QY	501	GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
DB	1675	GAGGACGAGTGTGGCGGAGGGCTGGCTGCCACAGCTGTGGCCCGGAGGCACTGC	1734
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
DB	1735	TGGGTGCCAGGGCCACCAGGTGTCAACTGCAGCCAGTTCTCTGGGGCCAGGAGTGC	1794
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
DB	1795	GTGGAGGAATGCGGAGTACTGCAGGGCTCCCGAGGGAGTATGTAATGCCAGGCACTGT	1854
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
DB	1855	TTGCGCTGCCACCTGAGTGTGAGCCCAAGATGGCTCAGTACCTGTTTGGACCGGAG	1914
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
DB	1915	GCTGACCACTGTGGCTGTGCCACTATAAGGACCTCCCTCTGCTGGCCCGCTGC	1974
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
DB	1975	CCCAGGGTGTGAACCTGACCTCTCTACATGCCCATGCCCATCTGGAAGTTTCCAGATGAGGAG	2034
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
DB	2035	GCGCATGCCACCTTGGCCCATCAACTGCACCCACTCCCTGTGTGGACCTGGATGACAAG	2094
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
DB	2095	GGCTGCCCGCGAGCAGAGAGCCCTCTGACGTCCATCATCTCTCGGGTGTGGC	2154
QY	653	-----	653
DB	2155	ATTCTGCTGGTGTGCTTTGGGGGTGCTTTGGGATCCTCATCAAGCGACGCGACGAG	2214

QY 653 ----- 653  
Db 2215 AAGATCCGGAAGTACACGATCGGAGACTGTCGAGAAACGGAGCTGGTGGAGCCGCTG 2274  
QY 653 ----- 653  
Db 2275 ACACCTAGCGGAGCGATGCCAACCCAGCGCGAGATGCGGATCCTGAAAGACGAGAGCTG 2334  
QY 653 ----- 653  
Db 2335 AGGAAGTGAAGGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2394  
QY 653 ----- 653  
Db 2395 CCTGATGGGAGAAATGTGAAAATTCACGTGGCCATCAAAGTGTGAGGGAAAAACACATCC 2454  
QY 653 ----- 653  
Db 2455 CCCAAGCCAAAGAAATCTTAGACGAAGCATACGTGATGGTGTGGGTCCCCA 2514  
QY 653 ----- 653  
Db 2515 TATGCTCCCGCTTCTGGGATCTGCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2574  
QY 653 ----- 653  
Db 2575 ATGCCCTATGGCTGCCCTTTAGACCATGTCCGGGAAACCGCGACGCCCTGGGCTCCAG 2634  
QY 653 ----- 653  
Db 2635 GACCTGTGAACGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG 2694  
QY 653 ----- 653  
Db 2695 CTCGTACACAGGACTTGGCGCTCGGAACGTGTGCTCAAGAGTCCCAACCATGTCAAA 2754  
QY 653 ----- 653  
Db 2755 ATTACAGACTTGGGCTGGCTCGGCTGTGGACATTTGACGAGACAGAGTACCATGCAGAT 2814  
QY 653 ----- 653  
Db 2815 GGGGGCAAGTGCCTCAAGTGGATGGCGTGGAGTCCATTTCTCCCGCGGCTTCCACC 2874  
QY 653 ----- 653  
Db 2875 CACCAGATGATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2934  
QY 653 ----- 653  
Db 2935 AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGTGGAAGGGGGGCGG 2994  
QY 653 ----- 653  
Db 2995 CTGCCCCAGCCCCCATCTGCACCAATGATGTCTACATGATCATGTTGAAATGTGGATG 3054  
QY 653 ----- 653  
Db 3055 ATTGACTCTGAATGTCCGCCAAGATTCCGGGAGTTGGTGTGTAATTTCTCCCGCATGGCC 3114  
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664  
Db 3115 AGGGACCCCGAGCGCTTTGTGGTTCATCCAGAAATGAGGACTTGGGCCCGCAGTCCCTTG 3174  
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684  
Db 3175 GACAGCACCTTCTACCGCTCACGTCTGGAGACGATGACATGGGGGACCTGGTGGATGCT 3234  
QY 685 GluGlyTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704  
Db 3235 GAGGAGTATCTGGTACCCCGACGAGGGCTTCTTGTGTCCACACCTGCCCCGGGCGCTGG 3294  
QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724

Db 3295 GGCATGTCTCCACCACAGGCACCGAGCTCATCTACAGGAGTGGCGTGGGACCTGCACA 3354  
QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744  
Db 3355 CTAGGCTGGAGCCCTCTGAAGAGAGAGCCCCAGGTCTCCACGTGGCACCTCCGAAGGG 3414  
QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764  
Db 3415 GCTGGCTCCGATGTATTTGATGTGACCTGGGAATGGGGCAGCAAGGGGCTGCAAGC 3474  
QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
Db 3475 CTCCCCACATGACCCCGCCCTCTACGGGTACAGTGGAGGACCCACACAGTACCCCTG 3534  
QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
Db 3535 CCCTCTGAGACTGATGGCTAGTTGGCCCCCTGACCTGCAGCCCCCAGCCTGAATATGTG 3594  
QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824  
Db 3595 AACCCAGCAGATGTTGGCCCCCAGCCCTTCGCCCGGAGAGGGCCCTCTGCTGCTGCC 3654  
QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
Db 3655 CGACCTGCTGTCCTGCTGGAAGGCCCAAGACTCTCTCCCGAGGGAAGATGGGTC 3714  
QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
Db 3715 GTCAAGACGCTTTTGGCGGTGGCGTGGAGAACCCCGAGTACTTTGACACCCAG 3774  
QY 865 GlyGlyAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884  
Db 3775 GGAGAGCTGCCCTCAGCCCCACCTCTCTGCTGCTTACGCCAGCCTTCGACACCTC 3834  
QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
Db 3835 TATTACTGGGACCGAGGACCCAGAGCGGGGCTCCACCCAGCAGCTTCAAGGACA 3894  
QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
Db 3895 CCTACGGCAGAAACCCAGAGTACCTGGGTCTGGACGTGCAGTG 3939

RESULT 5  
US-10-101-510-81  
; Sequence 81, Application US/10101510  
; Publication No. US20030148295A1  
; GENERAL INFORMATION:  
; APPLICANT: WAN, JACKSON  
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE  
; FILE REFERENCE: 15117, 0012  
; CURRENT APPLICATION NUMBER: US/10/101,510  
; CURRENT FILING DATE: 2002-03-20  
; PRIOR APPLICATION NUMBER: 60/276,947  
; PRIOR FILING DATE: 2001-03-20  
; NUMBER OF SEQ ID NOS: 805  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 81  
; LENGTH: 4473  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-101-510-81

Alignment Scores:  
Pred. No.: 0 Length: 4473  
Score: 4900.00 Matches: 919  
Percent Similarity: 73.23% Conservative: 0  
Best Local Similarity: 73.23% Mismatches: 0  
Query Match: 96.49% Indels: 336  
DB: 12 Gaps: 1

SEQ6 (1-919) x US-10-101-510-81 (1-4473)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20  
 DB 175 ATGGAGCTGGCGCTGTGGCGCTGGGGCTCTCTGCGCCCTCTGCGCCCGGAGCC 234  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 DB 235 GCGAGCACCAAGTGTGACCGGCACAGACATGAAGCTGGCTGCTGCCAGTCCCGAG 294  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
 DB 295 ACCCACTGGACATGCTCGCCACCTCTACAGAGGCTGCCAGGTGGTCAGGGAACCTG 354  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80  
 DB 355 GAACCTCACCTAGCTGCCACCAATGCCAGCTGTCTCTCCTCAGGATATCCAGAGGTG 414  
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 DB 415 CAGGGCTAGCTGCTCATCGCTCACAAACCAAGTGAGCGAGGTCCCACTGCAGAGGTGGG 474  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 DB 475 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGTGTACAAATGGA 534  
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 DB 535 GACCCGCTGAACAATACCACTCTGTTCACAGGGCTCCCAAGAGGCTCGCGGAGCTG 594  
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160  
 DB 595 CAGCTTCGAAGCTCACAGAGATCTTGAAGAGGGGTCTTATCAGCGGAACCCCAAG 654  
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180  
 DB 655 CTCTGCTACAGACACGATTTTGTGAAGGACATCTTCCAAAGAACCAACAGCTGGCT 714  
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 DB 715 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 774  
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 DB 775 GGTCCCGCTGTGGGAGAGAGTTCTGAGGATTTGACAGGCTCAGCGCACTGTCTGT 834  
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnLncys 240  
 DB 835 GCCGTGGCTGTGGCGCTGCAAGGGCCACTGCCCACTGCTGCTGCCATGAGCAGTGT 894  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 DB 895 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCCCTGCCCTCCACTTCAACCA 954  
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 DB 955 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTACAAACACACACACGTTTGA 1014  
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 DB 1015 TCCATGCCCAATCCCGAGGCGCGTATATATTGCGGCCAGCTGTGTGACTGCTGTGCC 1074  
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 DB 1075 TACAACCTACTTTTACGGACCTGGGATCTTCACCCCTGCTGCGCCCTGCACAAACAA 1134  
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 DB 1135 GAGGTACACGACAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGGAAGCCCTGTGCCCGA 1194  
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 DB 1195 GTGTGCTATGTGGCATGAGACACTTGCAGAGAGGTGAGGCGATTACCACTGCAAT 1254

QY 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 DB 1255 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGAGGCTGGCATTTCTGCGGAGAGC 1314  
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
 DB 1315 TTTGTATGGGAGCCAGCCCTCCAAACATGCTGCCCGCTCCAGCCAGAGCAGTCCAAGTGT 1374  
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420  
 DB 1375 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGGCCGACAGCCTGCCT 1434  
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 DB 1435 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGGAGCAATCTGCACAATGGCGC 1494  
 QY 441 TyrSerLeuThrLeuGlnGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 DB 1495 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGCTCACTGAGGGAA 1554  
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480  
 DB 1555 CTGGCAGTGGACTGGCCCTCATCCACCATACACCCACTCTGCTTCTGCACAGGGTG 1614  
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 DB 1615 CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACTGCCAACCGGCA 1674  
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 DB 1675 GAGGACGAGTGTGGCGAGGGCTGGCTGCCACCACTGTCGCCCGCAGGGCACTGC 1734  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 DB 1735 TGGGGTCCAGGCCCCCACCAGTGTGTCAACTGCAGCCAGTTCTTCTCGGGCCAGGAGTGC 1794  
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
 DB 1795 GTGGAGGAATCCGAGTACTGCAGGGGCTCCCAAGGAGTATGTGAATGCCAGGCACTGT 1854  
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 DB 1855 TTGCGCTGCCACCCCTGAGTGTGAGCCCAAGATGGCTCAGTACCTGTTTGGACCGGAG 1914  
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
 DB 1915 GCTGACCACTGTGGCTGTGCCCACTATAGGACCTCCCTTCTGCGCTGGCCCGCTGC 1974  
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
 DB 1975 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2034  
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640  
 DB 2035 GCGCATGTCAGCCCTTGGCCCATCAACTGCACCCACTCTGTTGAGACCTGGATGATCAAG 2094  
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
 DB 2095 GGCTGCCCGCCGACGACGACAGACCCCTCTTGAGCTCCATCATCTCTCGGGTGTGGC 2154  
 QY 653 653  
 DB 2155 ATTCTGCTGGTGTGCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGCAGCAG 2214  
 QY 653 653  
 DB 2215 AAGATCCGGAAGTACACGATGCGGAGCTGCTGCAGGAAACGAGCTGTGTGAGCGCGCTG 2274  
 QY 653 653  
 DB 2275 ACACCTAGCGGAGCGATGCCCAACAGCGCAGATGCGGATCCTGTAAAGACGAGGAGCTG 2334  
 QY 653 653

Db 2335 AGGAAGTGAAGTCTTGGATCTGGCGCTTTTGGCAGAGTCTACAAAGGGCATCTGGATC 2394  
QY 653 ----- 653  
Db 2395 CCTGATGGGAGAAATGTGAAAATTCAGTGGCCATCAAAGTGTGAGGGAAAAACACATCC 2454  
QY 653 ----- 653  
Db 2455 CCCAAGCCACAAAGAAATCTTAGACCAAGCATACGTGATGGCTGGTGGCTCCCA 2514  
QY 653 ----- 653  
Db 2515 TATGTCTCCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2574  
QY 653 ----- 653  
Db 2575 ATGCCCTATGGCTGCCTTAGACCATGTCCGGGAAACCCCGGAGCGCTGGGCTCCAG 2634  
QY 653 ----- 653  
Db 2635 GACCTGTGAATGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGGG 2694  
QY 653 ----- 653  
Db 2695 CTCGTACACAGGGACTTGGCGCTCGGAACGTGCTGTCAAGAGTCCCAACCATGTCAAA 2754  
QY 653 ----- 653  
Db 2755 ATTACAGACTCGGGCTGGCTGGCTGCTGGACATTTGACGAGACAGAGTACCATGCAGAT 2814  
QY 653 ----- 653  
Db 2815 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGCTTCACC 2874  
QY 653 ----- 653  
Db 2875 CACCAGATGATGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATCACTTTTGGGGCC 2934  
QY 653 ----- 653  
Db 2935 AAACCTTACGATGGATCCCAAGCCCGGGAGATCCCTGACCTGTGGAAAAGGGGAGCGG 2994  
QY 653 ----- 653  
Db 2995 CTGCCCGAGCCCCATCTGCACCATTTGATGTCTACATGATCATGGTCAAATGTTGGATG 3054  
QY 653 ----- 653  
Db 3055 ATTGACTCTGAATGTGGGCCAAGATTCCGGGAGTTGGTGTCTGAATTTCTCCCGCATGGCC 3114  
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664  
Db 3115 AGGACCCCGAGCCCTTGTGGTCTATCCAGATGAGGACTTGGCCCGCAGCAGTCCCTTG 3174  
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684  
Db 3175 GACAGCACCTTCTACCGCTCACTGCTGGAGGACCATGACATGGGGGACCTGGTGGATGCT 3234  
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704  
Db 3235 GAGGAGTATCTGGTACCCAGCAGGGCTTCTTCTCCAGACCTGCGCCGGCGCTGGG 3294  
QY 705 GlyMetValHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724  
Db 3295 GGCATGTCCACCACAGCCAGCCGAGCTCATCTACAGGAGTGGCGGTGGGGCCTGACA 3354  
QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744  
Db 3355 CTAGGGCTGGAGCCCTCTGAAGAGAGGAGCCCGCAGGCTCCACTGGCACCCTCCGAGGG 3414  
QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764  
|||||

Db 3415 GCTGGCTCCGATGATTTTGATGGTACCTGGGAATGGGGCAGCCAAAGGGCTGCAAGC 3474  
QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
Db 3475 CTCCCCACACATGACCCAGCCCTCTACAGGGGTACAGTGAGGACCCACACAGTACCCCTG 3534  
QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
Db 3535 CCCTCTGAGACTGATGGCTACGTTGCCCCCTGACCTGCAGCCCCCAGCCTGAATATGTG 3594  
QY 805 AsnGlnProAspValArgProGlnProSerProArgGluGlyProLeuProAlaAla 824  
Db 3595 AACCAAGCAGATGTTCGGCCCCAGCCCCCTTCGCCCGGAGAGGGCCCTCTGCCTGCTGCC 3654  
QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
Db 3655 CGACCTGCTGGTGCACACTCTGGAAGGCCCAAGACTCTCTCCCAAGGAAGANTGGGTG 3714  
QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
Db 3715 GTCAAAGACGTTTTCCTTGGGGTGCCTGGAGAACCCCGAGTACTTGACACCCAG 3774  
QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884  
Db 3775 GGAGGAGCTGCCCTCAGCCCCACCTCTCTGCTTTCAGCCCCAGCCTTCGACAACCTC 3834  
QY 885 TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
Db 3835 TATTACTGGGACAGGACCCACACAGAGGGGGGCTCCACCCAGCAGCCTTCANAGGGACA 3894  
QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
Db 3895 CCTACGGCAGAGAACCACAGACTACTCTGCTTTCAGCCCCAGCCTTCGACAGTG 3939

RESULT 6

US-10-146-473-32  
; Sequence 32, Application US/10146473  
; Publication No. US2003010888A1  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Matthew  
; APPLICANT: Gout, Ivan  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Gure, Ali  
; APPLICANT: Chen, Yao-Tsung  
; APPLICANT: Old, Lloyd  
; TITLE OF INVENTION: Breast Cancer Antigens  
; FILE REFERENCE: L00461/70130(JRV)  
; CURRENT APPLICATION NUMBER: US/10/146.473  
; PRIOR FILING DATE: 2002-05-15  
; PRIOR APPLICATION NUMBER: US 60/291,150  
; FILING DATE: 2001-05-15  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 32  
; LENGTH: 4473  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-146-473-32

Alignment Scores:  
Pred. No.: 0 Length: 4473  
Score: 4900.00 Matches: 919  
Percent Similarity: 73.23% Conservative: 0  
Best Local Similarity: 73.23% Mismatches: 0  
Query Match: 96.49% Indels: 336  
DB: 14 Gaps: 1

SEQ6 (1-919) x US-10-146-473-32 (1-4473)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20  
Db 175 ATGAGCTGGCGGCTTGTGCCCTGGGGGCTCTCTCCCTGCGCCCTTGTGCCCGCCGAGCC 234

QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
DB	235	GGAGACCAAGTGTGCACCGCACAGACATGAAGCTCGCGTCCCTGCCAGTCCCGAG	294
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60
DB	295	ACCCAGCTGGAGATGCTCCGCGACCTCTACCAAGGCTGCCAGGTGGTCAGGGAACCTG	354
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
DB	355	GAACCTACCTACTGCCACCAATGCCAGCTGTCTTCTCCAGGATATCCAGAGGTG	414
QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
DB	415	CAGGGCTACGTGCTCATCTCACAACCAAGTGAGCAGGTCCCAGTCGAGAGGTGCGG	474
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
DB	475	ATTGTGCGAGGACCCAGCTCTTTGAGGACAATATGCCCTGGCGCTGTAGACAATGGA	534
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
DB	535	GACCCGCTGAACAATACCAACCTGTGCAGAGGGCTCCCAAGAGGCTCGCGGAGCTG	594
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160
DB	595	CAGCTTCGAAGCTCACACAGATCTTTGAAGGAGGGGTCTTGATCCAGCGGAACCCAC	654
QY	161	LeuCysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnAsnGlnLeuAla	180
DB	655	CTCTGTACAGACACAGATTTTGTGAAGAGACATCTCCACAAGAACAACAGCTGGCT	714
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
DB	715	CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAAG	774
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
DB	775	GGCTCCCGTGTGGGAGAGAGTTCTGAGGATTTGTACAGAGCTGCAGCGCACTGTCTGT	834
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
DB	835	GCGGTGGCTGTGCGCGCTGCAAGGGGCCACTGCGCCACTGTCTGCCATGAGCAGTGT	894
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
DB	895	GCTGCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCGCTGCCCTCCACTTCAACCA	954
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
DB	955	AGTGGCATCTGTGAGCTGCACCTGCCAGCGCTGTGAGAGTGCAGCAACACAGACACG	1014
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
DB	1015	TCCATGCCCAATCCCGAGGGCGGTATACATTGCGGCGCAGCTGTGTGACTGCCCTGTCC	1074
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
DB	1075	TACAACCTACTTCTACGGACGTGGATCTGCACCCCTGCTGCCCCCTGCACACCAAA	1134
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
DB	1135	GAGGTGACACAGAGAGGATGGAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCGA	1194
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
DB	1195	GTGTGCTATGTGGCATGGAGCACTTGCAGAGGTGAGGCGAGTTACCACTGCAAT	1254
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
DB	1255	ATCCAGGAGTTGTGGCTGCAGAAGATCTTTGGGAGCCTGGCATTTCTGCGGAGAGC	1314
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe	400

DB	1315	TTTGTATGGGACCCAGCCTCCAAACAGTCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1374
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
DB	1375	GAGACTCTGGAAGAGATCACAGTTTACCTATACATCTCAGCATGGCCGACAGCCTGCCT	1434
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
DB	1435	GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGAGCAATTCTGCACAATGGGCC	1494
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
DB	1495	TACTCGCTCACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGCGCTCACTGAGGAA	1554
QY	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
DB	1555	CTGGCGAGTGGACTGGCCCTCATCCACCAATAACCCACCTCTGCTTCGCACACGGTG	1614
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
DB	1615	CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1674
QY	501	GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
DB	1675	GAGGACGAGTGTGGCGAGGGCTGGCTGCCACCACTGTGCGCCCGAGGCACTGC	1734
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
DB	1735	TGGGTGCCAGGCCCAACCCAGTGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGGATGC	1794
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
DB	1795	GTGGAGGAATCGCGAGTACTGCAGGGCTTCCCGAGGGAGTATGTGAATGCCAGGCACTGT	1854
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
DB	1855	TTGCGGTGCCACCTGAGTGTGAGCCCAAGATGGTCACTGAGTACCTGTTTGGACCGGAG	1914
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
DB	1915	GCTGACCATGTGTGGCCTGTGCCCACTATAGGACCCCTCCCTTCGTGCGGCCCGCTGC	1974
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
DB	1975	CCCAGCGGTGAAACCTGACCTCTCTTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	2034
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
DB	2035	GGCGCATGCGACCTTGCCCATCACTGCACCCACTCTGTGTGGACCTGGATGACAAG	2094
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer-	653
DB	2095	GGTCCCCCGGAGCAGAGAGCCGCTCTGACGCTCCATCATCTCTCGGTGGTGGC	2154
QY	653	-----	653
DB	2155	ATTCTGCTGGTGGTCTTTGGGGTGGTCTTTTGGGATCCTCATCAAGC3ACGGCAGCAG	2214
QY	653	-----	653
DB	2215	AAGATCCGGAAGTACACGATGGGAGCTGCTGCAGGAAACGAGGCTGGTGGAGCCGCTG	2274
QY	653	-----	653
DB	2275	ACACCTAGCGGAGCGATGCCCAACCGCCAGCGAGATCGCGATCCTGAAACAGACGAGCTG	2334
QY	653	-----	653
DB	2335	AGGAAGTGAAGTGTGGATCTGGCGCTTTTGGCACAGTCTTACAAGGGCATCTGGATC	2394
QY	653	-----	653



Db 2395 CCTGATGGGAGAATGTGAAATTCACAGTGGCCATCAAAGTGTGAGGGAAAAACACATCC 2454  
QY 653 ----- 653  
Db 2455 CCCAAAGCCAAAGAATCTTACAGCAACATACGTGATGGCTGGTGGCTCCCA 2514  
QY 653 ----- 653  
Db 2515 TATGTCTCCGCCCTTCTGGGCATCTGCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2574  
QY 653 ----- 653  
Db 2575 ATGCCCTATGGCTGCCTTTAGACCATGTCCGGGAAACCCGGAGCCCTGGGCTGCCAG 2634  
QY 653 ----- 653  
Db 2635 GACCTGCTGAACCTGGTGATGACATTGCCAAGGGGATGACCTACCTGGAGGATGTGGG 2694  
QY 653 ----- 653  
Db 2695 CTCGTACACAGGACCTTGGCCGCTCGGAACGTGCTGCTCAAGAGTCCCAACCATGTCAAA 2754  
QY 653 ----- 653  
Db 2755 ATTACAGACTTCGGGCTGGCTCGGCTGTGACATTTGACGACAGAGATACCATGCAGAT 2814  
QY 653 ----- 653  
Db 2815 GGGGCAAGGTGCCCATCAATGATGGATGGCGCTGGAGTCCATTCCTCCGCCGCGGTTCCAC 2874  
QY 653 ----- 653  
Db 2875 CACAGAGTGATGTGTGAGTTATGGTGTGATGCTGTGGGAGCTGATGACTTTTGGGGCC 2934  
QY 653 ----- 653  
Db 2935 AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAAGGGGAGCGG 2994  
QY 653 ----- 653  
Db 2995 CTGCCCCAGCCCCCATCTGCACCATGATGTCTACATGATCATGGTCAATGTGGATG 3054  
QY 653 ----- 653  
Db 3055 ATTGACTCTGAATGTCCGCCAAGATTCCGGGAGTTGGTCTGTAATTCCTCCGCATGGCC 3114  
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664  
Db 3115 AGGACCCCCAGCCCTTTGGTGCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTG 3174  
QY 665 AspSerThrPheTyrArgSerLeuGluAspAspMetGlyAspLeuValAspAla 684  
Db 3175 GACAGCACCTTCTACCGCTCACTCTGAGGACCATGACATGGGGAGCTGGTGGATGCT 3234  
QY 685 GluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGlyAlaGly 704  
Db 3235 GAGGAGTATCTGGTACCCAGCAGGGCTTCTCTCTCCAGACCTGCCCCGGGCGTGGG 3294  
QY 705 GlyMetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyAspLeuThr 724  
Db 3295 GGCATGTGTCCACACAGCCAGCCAGCTCAFTCTACAGGAGTGGCGGTGGGAGCTGACA 3354  
QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744  
Db 3355 CTAGGGCTGGAGCCCTCTGAGAGAGAGGCCCCAGGTCTCCACTGGCACCCCTCCAGAGG 3414  
QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764  
Db 3415 GCTGGCTCCGATGATTTGATGGTGACCTGGGAATGGGGCAGCAAGGGGCTGCCAAGC 3474  
QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
Db 3475 CTCCCCACATGACCCAGCCCTCTACAGGGGTACAGTGGAGGACCCACAGTACCCCTG 3534

QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
Db 3535 CCCTCTGAGACTGATGGCTAGCTTGGCCCCCTGACCTGCAGCCCCACGCTGAATATGTG 3594  
QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824  
Db 3595 AACGAGCCAGATGTTCGGCCCCAGCCCCCTTCGCCCCGAGAGGGGCCCTTCCTGCTGCTGCC 3654  
QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
Db 3655 CGACCTCTGTGGTGCACCTCTGGAAGGCCCAAGACTCTCTCCCAAGGAAATGGGGTCT 3714  
QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
Db 3715 GTCAAGACGCTTTTGGCTTTGGGGTGCCTGGAGAACCCCGAGTACTTGACACCCAG 3774  
QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884  
Db 3775 GGAGGAGTGGCCCTCAGCCCCACCTCTCTGCTTTCAGCCAGCCCTTCGACAACCTC 3834  
QY 885 TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
Db 3835 TATTACTGGGACAGGACCCACAGAGCGGGGGCTCCACCAGCAGCCTTCAAGGGACA 3894  
QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
Db 3895 CCTACGCGAGAGAACCCAGAGTACTGGTCTGGAGCTGCCAGTG 3939

RESULT 7

US-10-207-655-44  
; Sequence 44, Application US/10207655  
; Publication No. US20030118592A1  
; GENERAL INFORMATION:  
; APPLICANT: Ledbetter, Jeffrey A.  
; APPLICANT: Hayden-Ledbetter, Martha S.  
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
; FILE REFERENCE: 390069.401C1  
; CURRENT APPLICATION NUMBER: US/10/207,655  
; CURRENT FILING DATE: 2002-07-25  
; NUMBER OF SEQ ID NOS: 426  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 44  
; LENGTH: 4473  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-207-655-44

Alignment Scores:  
Pred. No.: 0 Length: 4473  
Score: 4900.00 Matches: 919  
Percent Similarity: 73.23% Conservative: 0  
Best Local Similarity: 73.23% Mismatches: 0  
Query Match: 96.49% Indels: 336  
DB: 14 Gaps: 1

SEQ6 (1-919) x US-10-207-655-44 (1-4473)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20  
Db 175 ATGGAGCTGGCGGCTTGTGCGCGTGGGGGTCTCTCTCGCCCTCTTGGCCCCGGAGCC 234  
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
Db 235 GCGAGCACCCAAAGTGTGCACCGGCACACATGAAGCTGGCGGCTCCCTTGCAGTCCCGAG 294  
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
Db 295 ACCACCTGGACATGCTCCGCCACCTCTACACGGGCTGCCAGGTGGTGACGGGAAACCTG 354  
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
Db 355 GAATCACCCTACCTGGCCCAATGCCAGCGCTGTCTTCTTCCAGGATATCCAGGAGGTG 414



```
Db 2575 ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAAACCGCGACGCCCTGGGCTCCAG 2634
Qy 653 ----- 653
Db 2635 GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG 2694
Qy 653 ----- 653
Db 2695 CTCGTACACAGGGACTTGGCGGCTCGGAACGTGCTGCTCAAGAGTCCCAACCATGTCAAA 2754
Qy 653 ----- 653
Db 2755 ATTACAGACTTGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGAGTACCATTGCAGAT 2814
Qy 653 ----- 653
Db 2815 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCCGCGGGTTCCACC 2874
Qy 653 ----- 653
Db 2875 CACCAGAGTGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2934
Qy 653 ----- 653
Db 2935 AAACCTTACGATGGATCCACGCGCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2994
Qy 653 ----- 653
Db 2995 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTGTCAAATGTTGGATG 3054
Qy 653 ----- 653
Db 3055 ATTGACTCTGAATGTCTGGCCCAAGATTCCGGGAGTTGCTGTGAATTTCTCCCGCATGGCC 3114
Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 3115 AGGGACCCCCAGCGCTTTGTGGTCAATCCAGAATGAGGACTTGGGCCCGCAGCTCCCTTG 3174
Qy 665 AspSerThrPheTyrArgSerLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3175 GACAGCACCTTCTACCGCTCCTGCTGGAGAGCATGACATGGGGGACCTGGTGGATGCT 3234
Qy 685 GluLeuTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3235 GAGGAGTATCTGTATCCACGAGGCGTCTTCTGTCCAGACCTGCCCGCGGCGTGGG 3294
Qy 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
Db 3295 GGCATGGTCCACACAGGCACCGCATCTATCCAGGAGTGGCGTGGGACCTGACA 3354
Qy 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3355 CTAGGGCTGAGGCGCTCTGAAGAGGAGGCCCCCAGGTCCTCACTGGCACCCCTCCGAAGG 3414
Qy 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeuGlnSer 764
Db 3415 GCTGGCTCCGATGTAITGATGTGACCTGGGAATGGGGCAGCCAGGGGCTGCRAAGC 3474
Qy 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db 3475 CTCCCCACATGACCCACGCCCTCTACAGCGGTACAGTGAGGACCCACAGTACCCTG 3534
Qy 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3535 CCCTCTGAGACTGATGGCTACGTTGCCCCCTGACCTGACGCCCCCAGCGCTGAATATGTG 3594
Qy 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3595 AACCAGCCAGATGTTGGCCCCCAGCCCCCTTCGCCCGGAGAGGCGCTCTGCTGCTGCC 3654
Qy 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
|||||
```

```
Db 3655 CGACCTGCTGGTGCCTCTTGGAAAGGCCCAAGACTCTCTCCCGAGGAAGAATGGGGTC 3714
Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
|||||
Db 3715 GTCAAGACGCTTTTGGCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCAG 3774
Qy 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884
|||||
Db 3775 GGAGGAGCTCCCTCAGCCCACTCCCTCCTGCTTCAAGCCAGCGCTTCGACAACCTC 3834
Qy 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
|||||
Db 3835 TATTACTGGGACGAGGACCCAGAGCGGGGGCTCCACCAGCACCTTCAAAAGGACA 3894
Qy 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
|||||
Db 3895 CCTACGGCAGAGAACCAGAGTACCTGGGTCTGGACGTGCCAGTG 3939

RESULT 8
US-10-207-498-5
; Sequence 5, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HEREULIN AND HER3
; FILE REFERENCE: 30448.103-US-01
; CURRENT APPLICATION NUMBER: US/10/207,498
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-207-498-5

Alignment Scores:
Pred. No.: 0 Length: 3765
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservative: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 12 Caps: 1

SEQ6 (1-919) x US-10-207-498-5 (1-3765)
Qy 1 MetGluLeuAlaAlaLeuCysArgTTPGlyLeuLeuAlaLeuLeuProGlyAla 20
|||||
Db 1 ATGGAGCTGGCGGCTTTGGCCCTGGGGCTCCTCTCGCCCTCTTGGCCCCCGGAGCC 60
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
|||||
Db 61 GCAGCACCAAGTGTGCACCGGCACAGACATGAAGTGGGCTCCCTGCCAGTCCCGAG 120
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
|||||
Db 121 ACCACCTGGACATGCTCCGCCACCTCTACACAGGGTGGCAGGTGGTGCAGGAAACCTG 180
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
|||||
Db 181 GAACCTCACCTACCTGCTGCCACCAATGCAGCTCTCTCTCTCAGGATATCCAGGAGGTG 240
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
|||||
```

Db 241 CAGGGCTAGCTGCTCATCGCTCACAAACCAGGTGAGCGAGTCCCTCCCTGAGGCTGGG 300  
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
Db 301 ATTGTGCGAGGACCCAGCTCTTTGAGGCACACTATGCCCTGGCCGTGTACAAATGGA 360  
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
Db 361 GACCCGCTGAACAATACCACTGCTTCACAGGGGCTCCCCAGGAGGCTCGGGAGCTG 420  
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160  
Db 421 CAGCTCGAAGCTCACAGAGATCTTGAAGAGAGGGTCTTATCATCAGCGGAACCCCA 480  
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180  
Db 481 CTCCTGCTACGACAGACAGATTTTGTGAAGAGACATCTTCCACAAGAACACAGCTGGCT 540  
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 600  
QY 201 GlySerArgCysTrpGlyLysSerGluAspCysGlnSerLeuThrArgThrValCys 220  
Db 601 GGCTCCCGCTGCTGGGAGAGAGTCTGTAGGATTTCTAGAGCCTGACGCGCACTGTCTGT 660  
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
Db 661 GCCGTGTGTGCCCCGTGCAAGGGCCACTGCCACTGACTGTGCTCCATGAGCAGTGT 720  
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
Db 721 GCTGCCGCTGACGGGCCCAAGCACTCTGACTGCTGCTGCCCTGCCCTCACTTCAACCCAC 780  
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
Db 781 AGTGGCATCTGTGAGCTGCCTGCCAGCCCTGGTCACTCAACACAGACAGTTTGAG 840  
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
Db 841 TCCATGCCCAATCCCGAGGGCCGTTATACATTCGGCGCCAGCTGTGTGACTGCCGTGCC 900  
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
Db 901 TACAACACTTCTTACGGACCTGGGATCTCTCACCCCTGCTGCCCCCTGCACAAACCAA 960  
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
Db 961 GAGGTGACACAGAGAGATGGAACACAGCCGTTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA 1020  
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
Db 1021 GTGTGCTATGCTGGGCATGGAGCACTTGCAGAGGTGAGGCGCAGTTACCACTGCCAAT 1080  
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
Db 1081 ATCCAGAGTTTGTGCTGGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCGGAGAGC 1140  
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe 400  
Db 1141 TTTGATGGGAGCCACGCTTCAACACTGCCCGCTCCAGCCAGAGCAGCTCAAGTGT 1200  
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420  
Db 1201 GAGACTCTGGAAGAGATCAGGTTACCTATACATCTCAGCATGGCGGACAGCCTGCCT 1260  
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
Db 1261 GACCTCAGGCTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTCGCACAATGGCGCC 1320  
QY 441 TyrSerLeuThrIleGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
Db 1321 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCACTGAGGGAA 1380

QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480  
Db 1381 CTGGGCAGTGGACTGGCCCTCATCCACCATAACCCACCTCTGCTTCTGTCACACGGTG 1440  
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
Db 1441 CCCTGGGACCACTTTTCGGAACCCGACCAAGCTCTGCTCCACTSCCAACGGCCA 1500  
QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
Db 1501 GAGGACGAGTGTGTGGCGAGGGCTGGCTGCCACCAGCTGTGCGCGAGGGCACTGC 1560  
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
Db 1561 TGGGGTCCAGGGCCCACTGCTGCTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1620  
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
Db 1621 GTGGAGGAATCCGAGTACTGCAAGGGCTCCCGAGGGAGTATGTGAATGCCAGGCACTGT 1680  
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
Db 1681 TTGCCGTGCCACCCCTGAGTGTCAAGCCCAAGATGGCTCAGTGCCTGTTTGGACCGGAG 1740  
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
Db 1741 GCTGACCACTGTGTGGCTGTGCCCTGTGCCACTATAGGACCCTCCCTTCTGCTGCCCGCTGC 1800  
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
Db 1801 CCCAGCGGTGTGAACCTGACCTCTCATATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860  
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640  
Db 1861 GCGCATGCCACCTTGGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGATGACAA 1920  
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
Db 1921 GGCTGCCCGCGAGCAGAGAGCCGCCCTCTGACGTCCATCTCTCGCGTGGTGGC 1980  
QY 653 653  
Db 1981 ATTCTGCTGCTGCTGCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAG 2040  
QY 653 653  
Db 2041 AAGATCCGGAAGTACACGATGCGGAGACTGTGCAGGAACGGAGCTGTGGAGCGCGCTG 2100  
QY 653 653  
Db 2101 ACACCTAGCGGAGCGATGCCCAACGAGCGCAGATGCGGATCCTGAAAGAGACGGAGCTG 2160  
QY 653 653  
Db 2161 AGGAAGGTGAAGTGTGCTGGGCTTTGGCACAGTCTTACAAAGGCATCTGGATC 2220  
QY 653 653  
Db 2221 CCTGATGGGGAATGTGAAAATTTCAGTGGCCATCAAAAGTTTGAGGGAACACATCC 2280  
QY 653 653  
Db 2281 CCCAAAGCCAAAGAAATCTTAGACGAAGCATACGTGATGCTGGTGTGGGCTCCCCA 2340  
QY 653 653  
Db 2341 TATGTCCTCCGCTTCTGGGCACTGCTGCCTGACATCCACGGTGCAGCTGTGTGACACAGCTT 2400  
QY 653 653  
Db 2401 ATGCCCTATGCTGCTCTTAGACCATGTCCGGGAAACCGCGGAGCCCTGGGCTCCCAG 2460

QY 653 ----- 653  
 Db 2461 GACCTGCTGAACCTGGTGTATGACAGATTCCCAAGGGGATGAGCTACCTGGAGATGTGCGG 2520  
 QY 653 ----- 653  
 Db 2521 CTCGTACACAGGGACTTGGCGGCTCGGAACGTGCTGTGTCAAGAGTCCCAACCATGTCAAA 2580  
 QY 653 ----- 653  
 Db 2581 ATTACAGACTTCGGGCTGGCTCGGCTGCTGAGCATTGACGAGACAGAGTACCATGCAGAT 2640  
 QY 653 ----- 653  
 Db 2641 GGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCCGCGCGGTTCAAC 2700  
 QY 653 ----- 653  
 Db 2701 CACCAGAGTGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2760  
 QY 653 ----- 653  
 Db 2761 AAACCTTACGATGGATCCCGAGCCCGGAGATCCCTGACCTGTGGAAGGGGGAGCGG 2820  
 QY 653 ----- 653  
 Db 2821 CTGCCCGAGCCCCCATCTGCACCATTTGATGCTTACATGATCATGTGTCAAATGTTGGATG 2880  
 QY 653 ----- 653  
 Db 2881 ATTGACTCTGAATGTCGGCCCAAGATTCGGGAGTGTGTGTAATTTCTCCCGCATGGCC 2940  
 QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664  
 Db 2941 AGGACCCCCAGCCCTTTGGTTCATCCAGATGAGGACTTGGCCCGAGCAGCTCCCTTG 3000  
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684  
 Db 3001 GACAGCACCTTCTACCCCTCCTGCTGAGGAGCATGACATGGGGGAGCTGGTGGATGCT 3060  
 QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704  
 Db 3061 GAGAGTATCTGGTACCCCGAGGAGGCTTCTGTCTCAGACCCCTGCCCGGGCGCTGG 3120  
 QY 705 GlyMetValHisAlaArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724  
 Db 3121 GGCATGCTCCACACAGCCAGCCAGCTCATCTACAGGAGTGGCGGTGGGGACCTGACA 3180  
 QY 725 LeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGluGly 744  
 Db 3181 CTAGGGCTGGAGCCCTCTGAAGAGAGGAGGCCCGCCAGGTCTCCACTGGCACCCCTCCGAAGG 3240  
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764  
 Db 3241 GCTGGCTCCGATGATTGATGGTGAAGTGGGAAATGGGGCAGCAGGAGGCTGCCAAGC 3300  
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
 Db 3301 CTCCCCACACATGACCCCGCCCTCTACAGCGGTACAGTACAGGACCCACACAGTACCCTG 3360  
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
 Db 3361 CCCTCTGAGACTGATGGCTACGTTGCCCTGACCTGACGCGCCCGACGCTGAATATGTG 3420  
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824  
 Db 3421 AACAGCCAGATGTTGGCCCCAGCCCCCTTCGCCCGAGAGGGCCCTCTGCTGCTGCC 3480  
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProGlyThrLeuSerProGlyLysAsnGlyVal 844  
 Db 3481 CGACCTGCTGGTCCACTCTGGAAAGGGCCCAAGACTCTCTCCCGAGGGAAGATGGGGTC 3540  
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValAlaGluAsnProGluTyrLeuThrProGln 864

Db 3541 GTCAAAAGACGCTTTTGGCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTTGACACCCAG 3600  
 QY 865 GlyGlyAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884  
 Db 3601 GGAGGACTGCCCTCAGCCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3660  
 QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
 Db 3661 TATTACTGGGACAGGACCCACAGAGCGGGGGCTCCACCCAGCACCTTCAAGGGGACA 3720  
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
 Db 3721 CCTACGCGAGAGAACCCAGAGTACCTGGGTCTGGAGCGTGCCAGTG 3765

RESULT 9  
 US-09-811-123-8  
 ; Sequence 8, Application US/09811123  
 ; Patent No. US20020001587A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sharon Erickson  
 ; APPLICANT: Mark Slikowski  
 ; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB  
 ; FILE REFERENCE: GENENT.073A2  
 ; CURRENT FILING DATE: 2001-03-16  
 ; PRIOR FILING DATE: 2001-03-16  
 ; PRIOR FILING DATE: 2000-10-05  
 ; PRIOR FILING DATE: 2000-06-23  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 8  
 ; LENGTH: 3768  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-811-123-8

Alignment Scores:  
 Pred. No.: 0 Length: 3768  
 Score: 4892.00 Matches: 918  
 Percent Similarity: 73.15% Conservative: 0  
 Best Local Similarity: 73.15% Mismatches: 1  
 Query Match: 96.34% Indels: 336  
 DB: 9 Gaps: 1

SEQ6 (1-919) x US-09-811-123-8 (1-3768)  
 QY 1 MetGluLeuAlaAlaLeuCysArgTTPGlyLeuLeuLeuAlaLeuLeuProGlyVala 20  
 Db 1 ATGAGCTGGCGGCTTGTGCTGGGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60  
 QY 21 AlaserThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 Db 61 GCGAGCACCAAGTGTGCACCGGCACACAGATGAAGCTGCGCTCCCTGCCAGTCCCGAG 120  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
 Db 121 ACCACCTGGACATGTCTCCGCCACCTCTACAGGGCTGCCAGGTGTGTGAGGAAACCTG 180  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
 Db 181 GAATCTACCTTACCTTGGCCCAATGCCAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 Db 241 CAGGGCTACGTGCTCATCGCTCACAAAGTAGGAGGAGGCTCCCACTGACAGGCTCGCG 300  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 Db 301 ATTTGTGCGAGGCCACCCAGCTCTTTTGTAGGACAACTATGCCCTGGCGCTGTAGACAATGGA 360





QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla	180	QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	481	CTCTGCTACCAAGACACAGATTTGTGGAAGGACATCTTCCACAAGAACACAGCTGGCT	540	Db	1561	TGGGGTCCAGGGCCACCCAGTGTGTCAACTGCAGCAGTTCTTTCGGGGCCAGAGTGC	1620
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200	QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	541	CTCACACTGATAGACCAACACGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAA	600	Db	1621	GTGGAGGAATCCGAGTACTGCAGGGGTCCCCAGGAGTATGTGAATGCCAGGCACGT	1680
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220	QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	601	GGCTCCCGTGTGGGAGAGAGTTCTGAGGATTTGTACAGGCTGACGGCCTGCTGT	660	Db	1681	TTGCCGTGCCACCTTGAGTGCAGCCCAAGAAATGGCTAGTACCTCTTTTGGACCGGAG	1740
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCys	240	QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	661	GCCGCTGGCTGTGCCGTGCAAGGGCCACCTGCTGCTGCCCTGCTCCACTTCAACCAC	720	Db	1741	GCTGACCAGTGTGGCCTGTGCCACATATAAGGACCTCCCTTCTGCTGGCCCGCTGC	1800
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260	QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	721	GCTGCCGTGCACGGGCCCAAGCACTCTGACTGCCTGCCCTGCTCCACTTCAACCAC	780	Db	1801	CCACGCGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280	QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640
Db	781	AGTGGCATCTGTAGCTGCACCTGCCAGCCCTGGTCACTACACACACACACGTTTGAG	840	Db	1861	GGCGCATGCCACCTTGCCCATCACTGCACCCACTCTCTGTGTGACCTGGATGACAAG	1920
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300	QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	841	TCCATGCCCAATCCCGAGGGCCGTATACATTCGGCGCCAGCTGTGTGACTGCCGTGCC	900	Db	1921	GGCTGCCCGCGAGCAGAGAGCCGCTCTGACGTCCATCGTCTCTCGCGTGGTGGC	1980
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320	QY	653	-----	653
Db	901	TACAACTACTTTTACGGACCTGGGATCTCTCACCCCTGCTGCCCCCTGCACAACCAA	960	Db	1981	ATTCTGCTGGTGTGGTCTTTGGGGTGGTCTTTGGGATCTCTCAAGGACGGCAGCAG	2040
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340	QY	653	-----	653
Db	961	GAGGTGACAGCAGAGATGGACACAGCGGTGTGAGAGTGACGAAGCCCTGTGCCCGA	1020	Db	2041	AAGATCCGGAAGTACAGATGCGGAGACTGTGTGAGGAAACGAGCTGTGGAGCGCGTG	2100
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	QY	653	-----	653
Db	1021	GTGTGCTATGTCTGGCATGAGCATTGCGAGAGGTGAGGCGAGTTACCAGTGCCCAAT	1080	Db	2101	ACACCTAGCGGAGCGATGCCAACCAAGCGGAGATGCGGATCCTGAAAGAGACGAGCTG	2160
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380	QY	653	-----	653
Db	1081	ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAG	1140	Db	2161	AGGAAGGTGAAGTGTCTTGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2220
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe	400	QY	653	-----	653
Db	1141	TTTGATGGGACCCAGCCTCCAACTGCCCCGCTCCAGCCAGACAGCTCCCAAGTGT	1200	Db	2221	CCTGATGGGGAATGTGAAAATTCACAGTGGCCATCAAAGTGTGAGGGAAAACACATCC	2280
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420	QY	653	-----	653
Db	1201	GAGACTCTGGAAGAGATCACAGTTTACCTATACATCTCAGCATGCGCGACAGCTGCCT	1260	Db	2281	CCCAAAGCCAACAAGAAATCTTAGACGAAGCATACGTGATGGTGTGGGTCCCCCA	2340
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440	QY	653	-----	653
Db	1261	GACCTCAGCGCTCTCCAGAACCTGCAGTAATCCGGGGAGCAATCTGCACAATGGCGCC	1320	Db	2341	TATGTCCTCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGTGTGACACAGCT	2400
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460	QY	653	-----	653
Db	1321	TACTCGCTACCTGCAAGGGCTGGCATCAGCTGGCTGGGGCTGGCTCCTGAGGGAA	1380	Db	2401	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAACCGGAGCGCTGGGCTCCCA	2460
QY	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480	QY	653	-----	653
Db	1381	CTGGGCACTGGAGCTGCGCTCTACCAATAACACCCACCTCTGCTTGTGTCACACGGTG	1440	Db	2461	GACCTGCTGAAGTGTGTATGCAGATGTGCCAAGGGGATGAGCTACTGTGGAGGATGTGCG	2520
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500	QY	653	-----	653
Db	1441	CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1500	Db	2521	CTCGTACACAGGAGTGTGGCGCTCGGAACGTGTGTGTCAGAGTCCCAACCATGTCAAA	2580
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520	QY	653	-----	653
Db	1501	GAGGAGAGTGTGGGCGAGGGCTTGGCTTCCACACAGCTGTGCGCCGAGGGCACTGC	1560	Db	2581	ATTACAGACTTCGGGCTGGCTCGGCTGTGGACATTGACGACACAGATTACCATGCAGAT	2640
				QY	653	-----	653





Db	751	GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTGTGAGAGCTGACGCGCACTGTCTGT	810	QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaAraCys	600
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240	Db	1891	GCTGACCACTGTGTGCGCTGTGCAAGGGCCACTGCCCCACTGACTGTGCCATGAGCAGTGT	1950
Db	811	GCGGTGGCTGTGCGCGCTGCAAGGGCCACTGCCCCACTGACTGTGCCATGAGCAGTGT	870	QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260	Db	1951	CCGAGGGTGTAAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	2010
Db	871	GCTGCCGGTGCACGGGCCCAAGCACTCTGACTGCTGCGCTGCCTCCACTTCAACCCAC	930	QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280	Db	2011	GGCGCATGCCAGCTTGGCCCCATCAACTGCACCCACTCTCTGTGTGACCTGGATGACAA	2070
Db	931	AGTGGCATGTGAGTGCACCTGCCAGCCCTGGTCACCTACAACACACACACGTTTGAG	990	QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300	Db	2071	GGCTGCCCGCGAGCAGAGCCGCCCTCTGACGTCCATCGTCTCTGCGGTGGTTGGC	2130
Db	991	TCCATGCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCCGTGCC	1050	QY	653	-----	653
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320	Db	2131	ATTCTGCTGCTGCTGTGCTTTGGGGTGGTCTTTTGGGATCCTCATCAAGGACGCGCAG	2190
Db	1051	TACAACCTAGCTTTCTACGACGTGGGATCTGCACCTGCTGCTGCCCTGCACAAACCAA	1110	QY	653	-----	653
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340	Db	2191	AAGATCCGAAAGTACACGATGCGGAGACTGCTGCAGAAACGAGAGCTGTGGAGCCGCTG	2250
Db	1111	GAGGTACACACGAGAGATGGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1170	QY	653	-----	653
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	Db	2251	ACACCTAGCGGAGCGATGCCAACACAGCGCAGATCGCGNCTCTGAAAGACGAGGACTG	2310
Db	1171	GTGTGCTATGGTCTGGGCATGAGCACTTGCAGAGGTGAGGCGACTTACCAGTGCCTAT	1230	QY	653	-----	653
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380	Db	2311	AGGAAGGTGAAGTGTGCTGCTGCGCTTTTGGCACAGTCTACAAGGCGCATCTGGATC	2370
Db	1231	ATCCAGAGTGTGCTGGCTGCAAGAAGATCTTTGGGAGCTGGCATTTCTGCCGAGAGC	1290	QY	653	-----	653
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe	400	Db	2371	CCTGATGGGAGAAATGTGAAATTCACGTGGCCATCAAAGTTGTGAGGGAACACATCC	2430
Db	1291	TTTGATGGGACCCAGCCCTCCAACTGCCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1350	QY	653	-----	653
QY	401	GluThrLeuGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420	Db	2431	CCAAAGCCAACAAGAAATCTTAGACGAAGCATACGTGATGCTGTTGGGCTCCCCA	2490
Db	1351	GAGACTCTGGAGAGATCAGAGTTTACCTATACATCTCAGCATGGCCGCGACAGCTGCCT	1410	QY	653	-----	653
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440	Db	2491	TATGTCCTCCGCTTCTGGGECATCTGCCCTGCATCCACGCTGCAGCTGTGTGACACAGCTT	2550
Db	1411	GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGGAGCAATTCGCACAAATGGCGC	1470	QY	653	-----	653
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460	Db	2551	ATGCCCTATGGCTGCCTCTTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCCAG	2610
Db	1471	TACTCCTGACCTGCACGGCTGGGCATCAGCTGGCTGGGGCTGCTCCTCAGTGGGAA	1530	QY	653	-----	653
QY	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480	Db	2611	GACCTGCTCAACTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG	2670
Db	1531	CTGGGAGTGGACTGGCCCTCCTCCACCATAACACCCACCTCTGCTTCTGTGCACACGGTG	1590	QY	653	-----	653
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500	Db	2671	CTCGTACACAGGACTTGGCCGCTCGGAACGCTGCTGGTCAAGAGTCCCCAACCATGTCAAA	2730
Db	1591	CCCTGGGACACGCTCTTTGGAAACCGCCACCAAGCTCTGCTCCACTGCCAACCGGCCA	1650	QY	653	-----	653
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520	Db	2731	ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGACACAGTACCATTGCAGAT	2790
Db	1651	GAGGACAGTGTGGGCGAGGCGCTGGCCCTGCCACAGCTGTGCGCCCGGCGGCACTGC	1710	QY	653	-----	653
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540	Db	2791	GGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTTCTCCGCGCGGCTTACC	2850
Db	1711	TGGGGTCCAGGGCCCAACCCAGTGTGTCAACTGCAGCCAGTTCTTTCGGGGCCAGGATGC	1770	QY	653	-----	653
QY	541	valGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560	Db	2851	CACCAAGATGATGTGTGGAGTTATGCTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2910
Db	1771	GTGGAGGAATTCGGAGTACTGCAAGGGGCTCCCGAGGGAGTATGTGAATGCCAGGCATGT	1830	QY	653	-----	653
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580	Db	2911	AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGTGAAAGGGGAGCGG	2970
Db	1831	TTGCCGTGCCACCTCAGTGTGACGCCCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG	1890				

```

Qy 653 ----- 653
Db 2971 CTGCCAGCCGCCCATCTGCACCATTTGATCTACATGATCATGTGCAATGTTGGATG 3030
Qy 653 ----- 653
Db 3031 ATTGACTCTGAATGTCGGCCAAAGATTCCGGAGTTGCTGTGAATTTCTCCCGCATGGCC 3090
Qy 654 ----- 654
Db 3091 AGGACCCCGCCGCTTTGGTTCATCCAGAAAGAGACTTGGCCCGACGAGTCCCTTG 3150
Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3151 GACAGCACCTTCTACCGCTCCTGAGAGACGATGACATGGGGACCTGGTGGATGCT 3210
Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3211 GAGGAGTATCTGTTACCCAGCAGCGAGGCTTCTTGTCCAGACCTGCCCCGGGCGCTGGG 3270
Qy 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724
Db 3271 GGATGTCTCACACAGCAGCGAGCTCATCTACAGAGAGTGGCGGTGGGACCTGACA 3330
Qy 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3331 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCGCCAGGTCTCCACTGGCACCTCCGAAGGG 3390
Qy 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db 3391 GCTGGCTCCGATGTTATTTGATGTGACCTGGGAATGGGAGCGAGCAAGGGGCTGCAAGC 3450
Qy 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db 3451 CTCCCCACATGACCCCGCCCTCTACAGCGGTACAGTGGAGGCCCGCCAGTACCCCTG 3510
Qy 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3511 CCCTCTGAGACTGATGCTAGCTGCTGCCCCCTGACCTGACGCCCGCCAGCTGATATGTG 3570
Qy 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3571 AACCAGCAGATGTTGCGCCCGCCAGCCCTCTGCGCCGAGAGGGGCCCTCTGCTGTGTC 3630
Qy 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3631 CGACCTGCTGGTCCACTCTGGAAGAGGCCAAGACTCTCTCCCGAGGGAAGAATGGGGTC 3690
Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3691 GTCAAGACGCTTTTGGCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCAG 3750
Qy 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3751 GGAGGAGCTGCCCTCAGCCCGCCAGCCCTCTCTGCTTCCAGCCAGCGCTTCGACAACTC 3810
Qy 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3811 TATTACTGGACGAGGAGGCCACAGAGCGGGGGCTCCACCCAGCACCTTCAAGAGGACA 3870
Qy 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3871 CCTACGGCAGAGAACCAGAGTACCTGGGTCTGGACGTGCCAGTG 3915

```

RESULT 12

```

US-10-007-926A-119
; Sequence 119, Application US/10007926A
; Publication No. US20030143539A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE

```

```

; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1346-R-00
; CURRENT APPLICATION NUMBER: US/10/007,926A
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 119
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: v-erb-b2 avian erythroblastic leukemia viral
; OTHER INFORMATION: oncogene homolog 2 (neuro/glioblastoma derived
; OTHER INFORMATION: oncogene homolog) (ERBB2) gene.
US-10-007-926A-119

Alignment Scores:
Pred. No.: 0 Length: 4530
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservative: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 12 Gaps: 1

SEQ6 (1-919) x US-10-007-926A-119 (1-4530)
Qy 1 MetGluLeuAlaAlaLeuCysArgTyrGlyLeuLeuLeuAlaLeuProGlyAla 20
Db 151 ATGAGCTGGCGGCTTGTGCCCTGGGGCTCTCTCGCCCTTGTGCCCGGAGCC 210
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 GCGAGCACCAAGTGTGCACCGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 270
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCACCTGGACATGCTCCGCCACCTCTACAGAGGCTGCCAGGTGGTGCAGGGAACCTG 330
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspLeuGluVal 80
Db 331 GAACCTACCTTACCTGCCCGCCCAATGCCAGCTGTCTCTCTCGAGGATATCCAGAGGTG 390
Qy 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTACGTGCTCATCGCTCACAAACAGTAGGCGAGGTCCCACTGCAGAGCTGGCG 450
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGGAGGACCCAGCTCTTTGAGGACAACCTATGCCCTGGCGCTGTAGACAATGA 510
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGCTGAACAAATACACCCCTGTCCAGGGGCCCTCCCGAGGAGGCTCGCGGAGCTG 570
Qy 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyGlyValLeuLeuGlnArgAsnProGln 160
Db 571 CAGCTCGAAGCCTCAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCGAG 630
Qy 161 LeuCystyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGTACCAGGACACGATTTTGTGAAGGACATCTTCCACAAGAAACACAGCTGGCT 690
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACACCAACCGCTCTCGGGGCTTCCAGCCCTGTCTTCCGATGTGTAAG 750
Qy 201 GlySerArgCysTyrGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220

```

D	b	751	GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATGTTCAGAGCCTGACGGCAGCTGTCTGT	810	Q	y	581	AlaAspGlnCysValalaCysAlaHisTyrIysAspProPheCysValalaArgCys	600
Q	y	221	AlaGlyGlyCysAlaArgCysGlyProLeuProThrAspCysCysHisGluGlnCys	240	D	b	1891	GCTGACCAGTGTGGCCTGTGCCACTATAAGGACCCCTCTCTGCTGCGCCGCTGC	1950
D	b	811	GCCGFTGGCTGTGCCGCTGCAAGGGCCACTGCCACTGACTGTGCCATGACAGTGT	870	Q	y	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Q	y	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260	D	b	1951	CCACGGGTGTGAACCTGACCTCTCATATGCCATCTGGAAGTTTCCAGATGAGGAG	2010
D	b	871	GCTGCCGCTGCACGGGCCCAAGCACTCTGACTGCTGCCCTGCCCTCCACTCAACAC	930	Q	y	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Q	y	261	SerGlyIleCysSluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280	D	b	2011	GCGCATGCCACCTTGCCCATCAACTGCACCCACTCTCTGTGGACCTGGATGACAAG	2070
D	b	931	ATGGCATCTGTAGCTGCACCTGCCAGCCCTGGTCACCTACACACACACACGTTGAG	990	Q	y	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Q	y	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300	D	b	2071	GGCTCCCGCGCAGCAGAGACCACCCCTCTGACGTCCATCGTCTCTCGGTGGTTGGC	2130
D	b	991	TCCATGCCCAATCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCGTGCC	1050	Q	y	653	-----	653
Q	y	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320	D	b	2131	ATTCTGCTGCTGCTGTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGCGCAGAG	2190
D	b	1051	TACAACCTACCTTTCTACGGACCTGGGATCTCTGCACCCCTGCTGCCCTGCACACCAA	1110	Q	y	653	-----	653
Q	y	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340	D	b	2191	AGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAACCGAGCTGCTGGAGCCGCTG	2250
D	b	1111	GAGGTGACACAGAGAGTGAACACACGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCCGA	1170	Q	y	653	-----	653
Q	y	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	D	b	2251	ACACCTAGCGGAGCGATGCCAACGAGCGCAGATGCGGATCCTGAAAGAGACGGAGCTG	2310
D	b	1171	GTGTGCTATGCTGGGCATGGACACTTGCAGAGGTGAGGGCAGTTACCAAGTGCCAT	1230	Q	y	653	-----	653
Q	y	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380	D	b	2311	AGGAAGTGAAGTGTCTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2370
D	b	1231	ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGGAGAGC	1290	Q	y	653	-----	653
Q	y	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe	400	D	b	2371	CCTGATGGGGAATGTGAAATTCAGTGGCCATCAAAGTGTGAGGAAACACATCC	2430
D	b	1291	TTTGATGGGACCCAGCCTCAACACTGCCCGCTCCAGCCAGACAGCTCCAAGTGT	1350	Q	y	653	-----	653
Q	y	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420	D	b	2431	CCCAAAGCAACAAGAAATCTTAGACGAAGCATACGTGCTGCTGTGGCTCCCCA	2490
D	b	1351	GAGACTCTGGAAGAGATCAGGTTACCTATACATCTCAGCATGCGCGACAGCCTGCCT	1410	Q	y	653	-----	653
Q	y	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440	D	b	2491	TATGTCTCCCGCTTCTGGGCATCTGCCCTGACATCCACGGTGCAGCTGTGTACACAGCTT	2550
D	b	1411	GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGAGCAATTCGCACAATGGCGCC	1470	Q	y	653	-----	653
Q	y	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460	D	b	2551	ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCGCGGACGCTGGCTCCCAG	2610
D	b	1471	TACTCGCTACCTGCAAGGGTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGGAA	1530	Q	y	653	-----	653
Q	y	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480	D	b	2611	GACCTGCTGAAGTGTGTATGCAGATTGCCAAGGGATGAGTACCTGGAGGATGTGCGG	2670
D	b	1531	CTGGGAGTGGACTGGCCCTCATCCACCATACACCCACTCTGCTTCTGTGCACACGGTG	1590	Q	y	653	-----	653
Q	y	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500	D	b	2671	CTCGTACACAGGAGCTTGGCCGCTCGGAACGTCTGTGTCAGAGTCTCCCAACCATGTCAA	2730
D	b	1591	CCCTGGGACAGCTCTTCCGAACCCGCGCACCAAGCTCTGCTCACACTGCCAACCGCCA	1650	Q	y	653	-----	653
Q	y	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520	D	b	2731	ATTACAGACTTCGGGCTGGCTCGGCTGTGGAGATTGACGAGACAGATACCATGCAGAT	2790
D	b	1651	GAGGACGAGTGTGGGCGAGGGCTTGGCTTCCACCACTGTGGCCCGGAGGCACTGC	1710	Q	y	653	-----	653
Q	y	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540	D	b	2791	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCGCCGCGCGGTTCCACC	2850
D	b	1711	TGGGGTCCAGGGCCACCCAGTGTGTCACTGCAGCCAGTTCTTTCGGGGCCAGGATGC	1770	Q	y	653	-----	653
Q	y	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560	D	b	2851	CACCAGAGTGTGTGGAGTTATGCTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2910
D	b	1771	GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCGACGGAGTATGTAAATGCCAGGCATGT	1830	Q	y	653	-----	653
Q	y	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580	D	b	2911	AAACCTTACGATGGGATCCACCGCGGAGATCCCTGACCTGCTGGAAAAGGGGGAGCGG	2970
D	b	1831	TTGCCGTGCCACCCCTGAGTGTACGCCCCAGAAATGGCTCAGTGACCTGTGTGGACCGGAG	1890					

```

QY 653 ----- 653
Db 2971 CTGCCCCAGCCCCCATCTGCACCATTTGATCTCTACATGATCATGGTCAAATGTTGGATG 3030
QY 653 ----- 653
Db 3031 ATTGACTCTGAATCTCGGCCAAGATTCGGGAGTGTGTCTGAATTCTCCCGCATGGCC 3090
QY 654 ----- 654
Db 3091 AGGGACCCCCAGCGCTTTGTGTCATCCAGATGAGGACTTGGGCCCGCAGCTCCCTTG 3150
QY 665 AspSerThrPheTyrArgSerLeuGluAspMetGlyAspLeuValAspAla 684
Db 3151 GACAGCCTTCTACCGCTCACTGCTGGAGACCATGACATGGGGAGCCTGGTGGATGCT 3210
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3211 GAGGAGTATCTGGTACCCAGCAGGGCTTCTTCTCCAGACCTGCCCGGGGCGCTGGG 3270
QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
Db 3271 GGCAATGTCCACACAGCAGCCAGCTCATCTACAGAGTGGCGGTGGGACCTGACA 3330
QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3331 CTAGGGCTGGAGCCCTCTGAAGAGAGAGGCCCGCCAGGTCTCCACTGGCACCCTCCGAAGG 3390
QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db 3391 GCTGGCTCCGATGATTGATGGTGACCTGGGAATGGGGCAGCAGAGGGCTGCAAGC 3450
QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db 3451 CTCCCCACACATGACCCCGCCCTCTACAGCGGTACAGTGAAGGCCACACAGTACCCTG 3510
QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3511 CCCCTCTGAGACTGATGGCTAGCTTGGCCCGCTGACCTGCAGCCCGCCAGCCTGAATATGTG 3570
QY 805 AsnGlnProAspValArgProGlnProSerProArgGluGlyProLeuProAlaAla 824
Db 3571 AACCAGCCAGATGTTGGCCCCCAGCCCCCTTCGCCCGCAGAGGGCCCTCTGCCCTGTGCC 3630
QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3631 CGACTCTGGTGGCACTCTGGAAGGGCCCAAGACTCTCTCCCGAGGAGAGATGGGGTC 3690
QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3691 GTCAAAGACGTTTTTGGCTTTGGGGTGCCGTGGAGAACCCCGAGTACTTTGACACCCAG 3750
QY 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3751 GGAGGAGTGGCTCCCTAGCCCCCAGCCCTCCCTCCCTCCCTCAGCCAGCCCTTCGACAACTC 3810
QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3811 TATTACTGGGACAGGACCCACAGAGGGGGGCTCCACCCAGCACCCTTCAAGAGGACA 3870
QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3871 CCTACGGCAGAGAACCCAGAGTACTCTGGGTCTGGACCTGCCAGTG 3915

```

RESULT 13

```

US-10-101-510-124
; Sequence 124, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012

```

```

; CURRENT APPLICATION NUMBER: US/10/101.510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 124
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-124

Alignment Scores:
Pred. No.: 0 Length: 4530
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservative: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 12 Gaps: 1

SEQ6 (1-919) x US-10-101-510-124 (1-4530)
QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCTCTCTCGCCCTCTTTCGCCCGGAGCC 210
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db* 211 GCAGACACCAAGTGTGCACCGCACAGACATGAAGCTGGCGCTTCTTCCAGGATATCCAGGAGTG 270
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu 60
Db 271 ACCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAACCTG 330
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspLeuGlnVal 80
Db 331 GAACCTACCTACCTGCGCCACCAATGCCAGCGCTGTCTTCTTCCAGGATATCCAGGAGTG 390
QY 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTACCTGCTCATCGCTCACACCAAGTAGGCGAGGTGCCACTGCAGAGAGCTCGG 450
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGGTGCTAGACAAATGA 510
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGCTGAACAATAACACCCCTGTGCACAGGGGCTTCCCGCAGGAGGCTCGCGGAGCTG 570
QY 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyValLeuLeuGlnArgAsnProGln 160
Db 571 CAGCTTGAAGACCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCGAG 630
QY 161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGCTACCCAGGACAGATTTTGTGGAAGACATCTTCCACAGAACAACACAGCTGGCT 690
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACACCAACCCGCTCTCGGGGCTGCCACCCCTGTCTTCCGATGTGTAAG 750
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCCGCTGCTGGGAGAGATTTCTGAGGATTTGTCAGAGCTTGACGCGCTGCTGTGT 810
QY 221 AlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db 811 GCCGGTGGCTGTCGCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 870
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 871 GCTGGCGGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGGCTGCCCTTCAACCCAC 930

```

261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 931 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACACACACAGACACGTTTGAG 990  
 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 991 TCCATGTCCCAATCCCGAGGCGCGGTATACATTTCGGCGCAGCTGTGTGACTGCCCTGTGCC 1050  
 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 1051 TACAACTACCTTTCACGAGCTGGGATCCTGCACCCCTGTCTGCCCTCGCACAAACCA 1110  
 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 1111 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA 1170  
 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 1171 GTGTGCTATGGTCTGGGCATGGAGCACTTGCAGAGGTGAGGGCACTTACCAGTGCCCAAT 1230  
 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 1231 ATCCAGGAGTTTGTGGCTGCAAGAGATCTTTGGAGCCTGGCATTTCTGCCGGAGAGC 1290  
 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe 400  
 1291 TTTGATGGGCGCCAGCCCTCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT 1350  
 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420  
 1351 GAGACTCTCGAAGAGATCAGAGTTACCTTATATCATCTCAGCATGGCGGACAGCCTGCCT 1410  
 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 1411 GACCTCAGCGCTTTCAGAACTTCGAACTGCAAGTAATCCGGGGAGCAATCTGCACAAATGCGCC 1470  
 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 1471 TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTGGCTGCTGCTGCTGCTGCTGCT 1530  
 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480  
 1531 CTGGGCGAGTGGAGCTGGCCCTCATCCACCAATACACCCACTCTGCTGCTGCTGCTGCTGCTGCT 1590  
 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 1591 CCTTGGGACCACTCTTTCGGAACCCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1650  
 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 1651 GAGGACGAGTGTGGGCGAGGCGCTGGCTGCCACCACTGTCGCCCGGAGGCGACATGC 1710  
 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 1711 TGGGGTCCAGGGCCCAACCACTGTGTCAACTGCAGCCAGTCTCTTCTGGGGCGCAGGAGTGC 1770  
 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
 1771 GTGGAGGAATGCGGAGTACTGCAGGGCTCCCGAGGAGTATGTGAATCCAGGCACTGT 1830  
 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 1831 TTGCCGTGCCACCTGAGTGTGAGCCCAAGATGGCTCAGTCAGCTGTTTGTGGACCGGAG 1890  
 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
 1891 GCTGACAGTGTGTGGCCCTGTCACCACTATAAGGACCCCTCCCTTCTGCGTGGCGCGCTGC 1950  
 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
 1951 CCCAGCGGTGTGAACCTGACCTCTCTCATATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010

QY GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640  
 Db GCGCATGCCAGCCTTGCCCATCAACTCACTGACCCCACTCTGTGTGACCTGATGACAAG 2070  
 QY GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653  
 Db GGCTCCCCCGCGAGCAGAGAGCCAGCCCTCTGACGTCCATCGTCTCTGCGGTGTGTGGC 2130  
 QY ----- 653  
 Db ATCTCTGCTGGTGGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAG 2190  
 QY ----- 653  
 Db AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGTGTGAGCGCGCTG 2250  
 QY ----- 653  
 Db ACACCTAGCGGAGCGATGCCCAACCGCGCAGATGCGGATCCTGAAAGAGACGGAGCTG 2310  
 QY ----- 653  
 Db AGAAGGTGAAGTCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2370  
 QY ----- 653  
 Db CCTGATGGGAGAATGTGAAAATTCAGTGGCCATCAAAAGTCTTGAGGGAACACACATCC 2430  
 QY ----- 653  
 Db CCCAAAGCCAAAGAAATCTTAGACGAAGCATACGTGATGGTGGTGTGGGCTCCCA 2490  
 QY ----- 653  
 Db TATGTCTCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGTGACACAGCTT 2550  
 QY ----- 653  
 Db ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCAG 2610  
 QY ----- 653  
 Db GACCTGCTGAATGCTGTATGCAGATTGCCAAGGGGATGAGTACCTGAGGAGTGTGGC 2670  
 QY ----- 653  
 Db CTCGTACACAGGAGCTTGGCCCTCGGAACGTGTGGTCAAGAGTCCCACCATGTCAAA 2730  
 QY ----- 653  
 Db ATTACAGACTTCGGCTGGCTCGGCTGGTGGATTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2790  
 QY ----- 653  
 Db GGGGCAAGTGCCTCAAGTGGATGGCGTGGAGTCCATTCCTCCGCGCGGCTTACCC 2850  
 QY ----- 653  
 Db CACCAGACTGATCTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2910  
 QY ----- 653  
 Db AAACCTTAGATGGATCCAGCCCGGAGATCCCTGACCTGCTGTGAAAAGGGGAGCGG 2970  
 QY ----- 653  
 Db CTGCCCCAGCCCCCATCTGCACCATGATGATGCTATGATCATGATCATGTTAAATGTTGGATG 3030  
 QY ----- 653  
 Db ATTGACTCTGAATTCGGCCAAAGATTCCGGGAGTTGGTGTCTGAATTCCTCCCGCATGGCC 3090  
 QY -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664

:	ORGANISM:	Homo sapiens
:	FEATURE:	
:	NAME/KEY:	CDS
:	LOCATION:	(151)..(3915)
:	OTHER INFORMATION:	
US-10-338-730-1		
Alignment Scores:		
Pred. No.:	0	Length: 4530
Score:	4892.00	Matches: 918
Percent Similarity:	73.15%	Conservative: 0
Best Local Similarity:	73.15%	Mismatches: 1
Query Match:	96.34%	Indels: 336
DB:	12	Gaps: 1
SEQ6 (1-919) x US-10-338-730-1 (1-4530)		
Qy	1	MetGluLeuAlaLaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla 20
Db	151	ATGGAGTGGGGCGCTTGCCGTGGGGCTCCTCCTGCCTTCGCCCTCGCCGCCGAGGC 210
Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db	211	GCAGACACCAAGTGTGCACCGGCACAGACATGAAGCTCGCGCTCCCTGCCAGTCCCAG 270
Qy	41	ThrHisLeuaspMetLeuAtrqHisLeuTyrrGlnGlyCysGlnValValcInGlyAsnLeu 60
Db	271	ACCACCTGGACATGCTCCGCCACTCTPACCAAGGCTGCCAGTGGTGCAGGAACCTG 330
Qy	61	GluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspileGlnGluVal 80
Db	331	GAACTCACCTACCTGCCACCAATGCCAGCTGTCTCTTCTGCAGGATATCCAGAGGTG 390
Qy	81	GlnGlyTyrrValLeulleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuAtrg 100
Db	391	CAGGGCTACGTGCTCATCGCTCAACAACAGTAGGAGGTCCTCCACTGCAGAGGCTGCGG 450
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAspAsnGly 120
Db	451	ATTGTGCGAGGCACCCAGCTCTTTGAGGCACAACATATGCCCTGGCCGTGCTACACAATGA 510
Qy	121	AspProLeuAsnAnsthrThrProvalThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db	511	GACCCGTGAACAAATACACCCCTGTCTCAGSGGCCCTCCCAGAGAGGCTGCGGGAGCTG 570
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db	571	CAGTTTCGAGCCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCACGGGAACCCCCAG 630
Qy	161	LeuCystTyrrGlnAspThrIleLeuTrpLysAspilePheHisLysAsnAsnGlnLeuAla 180
Db	631	CTCTGCTACCAAGGACAGATTTTGTGAAGGACATCTTCCACAAGAACAACCAAGCTGGCT 690
Qy	181	LeuThrLeuilleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db	691	CTCACACTGATAGACACCAACCGCTTCGGGGCCCTGCCACCCCTGTFTCTCCGATGTGAAG 750
Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgthrValCys 220
Db	751	GGCTCCCGCTGCTGGGGAGAGATTCTCAGGATTTCTCAGAGCCTGACCGGCACCTGCTGT 810
Qy	221	AlaGlyCysAlaArCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db	811	GC CGGTGGCTGTGCCCGCTGCAAGGGGGCACTGCGCCACTGCTGCTGCGCATGAGCAGTGT 870
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheasnhis 260
Db	871	GC TCGCGGCTGCACGGGCCCAACACTCTGACTGCTGGCTGGCTCCACTTCAACAC 930
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThrPheGlu 280
Db	931	AGTGGCATCTGTGAGCTGC ACTGCCAGCCCTGGTGCCTTACACACAGACACGCTTTTCAG 990

691	CTCACACTGATAGACACCAACCGCTCTGGGCGCTGCCACCCCTGTTCTCCGATGTGAAG	750
201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
751	GGCTCCCGCTGCTGGGAGAGAGTTCTGAGCATTTCTCAGAGCCTCAGCGCACTCTCTGT	810
221	AlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluClnCys	240
811	GC CGTGGCTGTGCCCGCTGCAAGGGCCACTGCCACTGCTGCTCATGAGCAGTGT	870
241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
871	GCTGCCGGCTGCACGGGCCCAACACTCTCACTGCCTGGCGCTGCCTCACTTCAACAC	930
261	SerGlyTleCysGluLeuHisCysProAlaLeuValThrTyAsnThrAspThrPheGlu	280
931	AGTGGCATCTGTGAGCTGCATGCTGCCAGCCCTGGTCACTACACACAGACACCGTTTCAG	990

```

RESULT 14
US-10/338-730-1
: Sequence 1, Application US/10338730
: Publication No. US20030147905A1
: GENERAL INFORMATION:
: APPLICANT: Genzyme Corporation
: APPLICANT: Nicolette, Charles A.
: TITLE OF INVENTION: THERAPEUTIC COMPOUNDS
: FILE REFERENCE: 5017C
: CURRENT APPLICATION NUMBER: US/10/338,730
: CURRENT FILING DATE: 2003-01-08
: PRIOR APPLICATION NUMBER: US 09/527,487
: PRIOR FILING DATE: 2002-03-16
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 4530
: TYPE: DNA

```







Dd	691	CTCACACTGATAGACACCAACCGCTCTCGGGCCCTGCCACCCCTGTTCTCCGATGTGTAAG	750
Qy	201	GlySerArgCysTrpGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Dd	751	GGCTCCCGCTGCTGGGAGAGAGTCTGAGGATGTTCAGAGCGCTGACGCGCACTGTCTGT	810
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Dd	811	GCCGGTGGCTGTGCCCGCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT	870
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Dd	871	GCTGCCGGCTGCACGGGGCCCAAGCACTCTGACTGCTGGCCCTGCCCTCAACCCAC	930
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Dd	931	AGTGGCATCTGTGAGCTGCACTGCCAGGCCCTGGTCACCTACACACAGACAGCTTTGAG	990
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Dd	991	TCCATGCCAAATCCCGAGGGCCGGTATACATTGGCGCCAGCTGTGTGACTGCCTGTCCC	1050
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Dd	1051	TACAACTACCTTTCTACGACGTGGGATCTTCACCCCTGCTGCCCCCTGCACAAACCAA	1110
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Dd	1111	GAGTCACAGCAGAGATGGAACACAGCGGTGTGAGAGTGCACGAAGCCCTGTGCCGA	1170
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Dd	1171	GTGTGTATGCTGGCATGGAGCACTTCCGAGAGGTGAGGCGCAATTACCACTGCCAAT	1230
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Dd	1231	ATCCAGAGTTTCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCAATTTCTGCCGGAGAGC	1290
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe	400
Dd	1291	TTTGTATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1350
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Dd	1351	GAAGACTTGGAGAGATCAAGTTTACCTATACATCTCAGCATGGCGGACAGCCTGCCT	1410
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Dd	1411	GACCTCAGCGTCTCCAGAACCTGCAAGTAATCCGGGGAGGAATTCGCACAAATGGCGCC	1470
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Dd	1471	TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAA	1530
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Dd	1531	CTGGGAGAGTGGCTGCTCACTCCACCACTCAACACCCCACTCTGCTTCGTGCACACGGTG	1590
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Dd	1591	CCCTGGGACCAAGCTCTTTCCGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1650
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Dd	1651	GAGGAGAGTGTGTGGCGAGGGCCTGGCCTGCCACAGCTGTGCCCGGAGGGCACTGC	1710
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Dd	1711	TGGGGTCCAGGGCCCACTGCTGCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC	1770
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Dd	1771	GTGGAGGAATGCCAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	1830
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Dd	1831	TTGCCGTGCCACCTGAGTGTGAGCCCAAGATGGCTCAGTACCTGTGTTTGGACCGGAG	1890
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Dd	1891	GCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCCTCCCTTCTGCTGCGCCGCTGC	1950
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Dd	1951	CCACGGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	2010
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Dd	2011	GGGCATGCCAGCCTTGCCCCCACTCACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	2070
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Dd	2071	GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACCTCCATCGTCTCTGCGGTGGTGGC	2130
Qy	653	-----	653
Dd	2131	ATTCTGCTGCTGCTGCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGCGACGAG	2190
Qy	653	-----	653
Dd	2191	AGATCCGGAAGTACAGATGCGGACACTGCTCCAGAAACGGAGCTGCTGGAGCCGCTG	2250
Qy	653	-----	653
Dd	2251	ACACCTAGCGGAGCGATGCCAACAGCGCAGATCGCGGATCCTGAAAGAGACGGAGCTG	2310
Qy	653	-----	653
Dd	2311	AGGAAGTGAAGTGTGATGCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2370
Qy	653	-----	653
Dd	2371	CCTGATGGGAGAAATGTGAAATTCAGTCAGTGCCATCAAAGTGTGAGGGAAACACATCC	2430
Qy	653	-----	653
Dd	2431	CCCAAAGCCAACAAAAATCTTAGACGAAGCATAGTGATGGCTGGTGTGGCTCCCCA	2490
Qy	653	-----	653
Dd	2491	TATGTCTCCGCTTCTGGGCATCTGCTTGACATCCACGGTGCAGCTGGTGACACAGCTT	2550
Qy	653	-----	653
Dd	2551	ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCCGGGACGCTGGGCTCCAG	2610
Qy	653	-----	653
Dd	2611	GACCTGCTGAAGTGTATGAGATTCGCAAGGGGATGAGTACCTGAGGATGTGCGG	2670
Qy	653	-----	653
Dd	2671	CTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAA	2730
Qy	653	-----	653
Dd	2731	ATTACAGACTTCGGGCTGGCTGGCTGGACATTGACGAGACAGATACCATGCAGAT	2790
Qy	653	-----	653
Dd	2791	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGCTTACC	2850
Qy	653	-----	653
Dd	2851	CACCAGTGTGTGTGGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2910

QY	653	-----	653
Db	2911	AAACCTTACGATGGGATCCAGCCGGGAGATCCTGACCTGCTGGAAGGGGAGCGG	2970
QY	653	-----	653
Db	2971	CTGCCCCAGCCCCCATCTGCACCATTTGATGCTACATGATCATGGTCAAATGTTGGATG	3030
QY	653	-----	653
Db	3031	ATTGACTCTGAATGTCGCCAAGATTCCGGGAGTTGGTGTCTGAATTTCTCCCGCATGGCC	3090
QY	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	3091	AGGGACCCCCAGCGCTTTGTGTCATCCAGAATGAGGACTTGGGCCAGCCAGCTCCCTTG	3150
QY	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3151	GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGGACCTGGTGGATGCT	3210
QY	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyValaGly	704
Db	3211	GAGGAGTATCTGGTACCCAGCAGGGCTTCTTCTCTCCAGACCTGCCCGGGCGCTGGG	3270
QY	705	GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr	724
Db	3271	GGCATGGTCCACACAGCAGCCGCGCTCATCTACAGGAGTGGCGGTGGGACCTGACA	3330
QY	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
Db	3331	CTAGGGCTGGAGCCCTCTGAAGAGAGGCCGCCAGGCTCCACTGGCACCTCCGAAGGG	3390
QY	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer	764
Db	3391	GCTGGCTCCGATGATTTGATGGTGACCTGGGAATGGGGGAGCAAGGGGCTGCAAAAGC	3450
QY	765	LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu	784
Db	3451	CTCCCCACACATGCCCCAGCCCTCTACAGCGGTACAGTGAGGAGGCCCCACAGTACCCCTG	3510
QY	785	ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal	804
Db	3511	CCCTCTGAGACTGATGGTACGTTGCCCCCTGACCTGCAGCCCCAGCCCTGAATATGTG	3570
QY	805	AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla	824
Db	3571	ANCCAGCCAGATGTTGGCCCCCAGCCCTTCGCCCGAGAGGGCCCTCTGCCCTGCTGCC	3630
QY	825	ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal	844
Db	3631	CGACTGTGTGTCACCTCTGGAAGGGCCCAAGACTCTCTCCCGAGGGAAGATGGGGTC	3690
QY	845	VallYsAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln	864
Db	3691	GTCAAAAGACGTTTGTGCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACACCCAG	3750
QY	865	GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu	884
Db	3751	GGAGGAGTGCCCTTCAGCCCCCAGCCCTCTCTGCTTCAGCCCGAGCTTCGACAACTC	3810
QY	885	TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr	904
Db	3811	TATTACTGGGACGAGGACCCACAGAGGGGGGCTCCACCCAGCACCTTCAAAGGGACA	3870
QY	905	ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal	919
Db	3871	CCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG	3915

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 22:55:19 : Search time 232.963 Seconds  
(without alignments)  
1741.185 Million cell updates/sec

Title: SE06

Perfect score: 5078

Sequence: 1 MELAALCRWGLLLALLPPGA.....TFKGTPTAENPEYLGLDVPV 919

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O/cgn2\_1/USPRO.spool/p/HOLLERAN80/runat\_15102003\_131913\_20570/app\_query.fasta\_1.4685  
-DB=Issued\_Patents\_NA -Qfmt=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN80 -CGN\_1\_1\_403@runat\_15102003\_131913\_20570 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents.NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4900	96.5	3768	2	US-08-625-101-1
2	4900	96.5	3768	2	US-08-356-786-1
3	4900	96.5	4473	2	US-09-048-804-1
4	4900	96.5	4473	3	US-09-056-105-26
5	4892	96.3	4530	1	US-08-229-515A-9
6	4892	96.3	4530	1	US-08-645-865-9
7	4892	96.3	4530	4	US-09-167-322-4
8	4892	96.3	4530	4	US-09-527-487-1
9	4892	96.3	4530	4	US-09-877-177A-11
10	4134	81.4	3955	1	US-08-229-515A-14
11	4134	81.4	3955	1	US-08-645-865-14
12	3632	71.5	2385	2	US-09-146-283-3

13	3632	71.5	2385	3	US-08-579-823A-3	Sequence 3, Appli
14	3632	71.5	2385	3	US-09-344-195-3	Sequence 3, Appli
15	3471	68.4	1872	3	US-08-422-108-2	Sequence 2, Appli
16	3471	68.4	1872	4	US-08-422-734-2	Sequence 2, Appli
17	1645	32.4	5532	2	US-08-475-035-3	Sequence 3, Appli
18	1645	32.4	5532	4	US-09-676-610B-17	Sequence 17, Appli
19	1528.5	30.1	5484	3	US-09-632-580A-3	Sequence 3, Appli
20	1528.5	30.1	5501	1	US-08-484-438-1	Sequence 1, Appli
21	1451	28.6	5555	1	US-08-484-438-3	Sequence 3, Appli
22	1435	28.3	4905	1	US-07-978-895-3	Sequence 3, Appli
23	1435	28.3	4905	1	US-08-473-119-3	Sequence 3, Appli
24	1435	28.3	4905	2	US-08-475-352-3	Sequence 3, Appli
25	1435	28.3	4975	3	US-09-630-706-3	Sequence 3, Appli
26	1311	25.8	4545	6	5183884-3	Patent No. 5183884
27	1032	20.3	1958	4	US-09-570-454-1	Sequence 1, Appli
28	1032	20.3	1958	4	US-09-867-521-1	Sequence 1, Appli
29	942	18.6	1593	4	US-09-676-610B-25	Sequence 25, Appli
30	942	18.6	1868	1	US-08-658-883B-1	Sequence 1, Appli
31	942	18.6	1868	4	US-09-676-610B-26	Sequence 26, Appli
32	493	9.7	322	1	US-08-421-356-1	Sequence 1, Appli
33	493	9.7	322	4	US-09-046-783-1	Sequence 1, Appli
34	381	7.5	4149	2	US-08-737-715-1	Sequence 1, Appli
35	376	7.4	2533	3	US-09-149-922-5	Sequence 5, Appli
36	333	6.6	4989	2	US-08-625-819-1	Sequence 1, Appli
37	333	6.6	4975	2	US-08-249-687C-1	Sequence 1, Appli
38	333	6.6	4989	2	US-08-666-392A-3	Sequence 3, Appli
39	333	6.6	4989	3	US-08-755-558-4	Sequence 4, Appli
40	333	6.6	4989	3	US-08-746-559A-1	Sequence 1, Appli
41	333	6.6	4989	3	US-08-880-313A-9	Sequence 9, Appli
42	333	6.6	4989	3	US-09-199-926-3	Sequence 3, Appli
43	333	6.6	4989	4	US-09-389-855A-9	Sequence 9, Appli
44	333	6.6	4989	4	US-09-668-822-9	Sequence 9, Appli
45	328	6.5	4993	3	US-08-746-559A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-08-625-101-1

; Sequence 1, Application US/08625101

; Patent No. 5869445

; GENERAL INFORMATION:

; APPLICANT: Cheever, Martin A.

; APPLICANT: Disis, Mary L.

; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE

; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION

; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu

; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/625,101

; FILING DATE: 01-APR-1996

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Sharkey, Richard G.

; REGISTRATION NUMBER: 32,629

; REFERENCE/DOCKET NUMBER: 920010.448C7

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3768 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..3765  
 ; US-08-625-101-1

## Alignment Scores:

Pred. No.: 0 Length: 3768  
 Score: 4900.00 Matches: 919  
 Percent Similarity: 73.23% Conservative: 0  
 Best Local Similarity: 73.23% Mismatches: 0  
 Query Match: 96.49% Indels: 336  
 DB: 2 Gaps: 1

## SE06 (1-919) x US-08-625-101-1 (1-3768)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20  
 DB 1 ATGGAGCTGGCGCCTGTGGCGCTGGGGGCTCCTCCTCGCCCTCTTGCCCCCGGAGCC 60  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 DB 61 GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGCTCGCGCTCCCTGCCAGTCCCGAG 120  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
 DB 121 ACCACCTGGACATGTCCGCCACCTCTACCAAGGCTGCCAGGTGGTGAGGGAACCTG 180  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
 DB 181 GAACTCACCTACCTGCCACCAATGCCAGCCTGTCTCTCGAGGATATCAGGAGGTG 240  
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 DB 241 CAGGGCTAGCTGTCTATCGCTCACACCAAGTGAGGAGGTGCCACTGCCAGAGGTCCGG 300  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 DB 301 ATTGTCCGAGGACCCAGCTCTTTGAGGACACTATGCCCTGGCGGTGTAGCAATGGA 360  
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 DB 361 GACCCCTGAACAATACCACCTGTCTCAGGGGCTCCCGAGGAGGCTCGGGAGCTG 420  
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160  
 DB 421 CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCAG 480  
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180  
 DB 481 CTCGTCTACGAGACAGATTGTGGAAGGACATCTCCACAAGAACCAACAGGTGGCT 540  
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 DB 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTAG 600  
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 DB 601 GGCTCCCGTGTGGGAGAGAGATTCTGAGGATTGTGAGAGCTGTCAGAGCTGACGGCCTGTCTGT 660  
 QY 221 AlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240  
 DB 661 GCGGTGGTGTGCGCGCTGCAAGGGGCCACTGCCCACTGTCTGCTGCTGCTGCTGCTGCTGCT 720  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 DB 721 GCTGCGGCTGACGCGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280

DB 781 ACTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACCAACACACACACGTTTGAG 840  
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 DB 841 TCCATGCCCATCCCGAGGCGCGGTATACATTCGGGCGCAGCTGTGTACTGCTGTCTCC 900  
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 DB 901 TACAACCTACCTTCTACGGACGTGGGATCCTGCACCTCGTCTGCCCTGTCACAAACCAA 960  
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 DB 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGAGCAAGCCCTGTGCCCCGA 1020  
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 DB 1021 GTGTGCTATGTGCTGGGATGGAGCACCTTGGCAGAGGTGAGGCGAGTTACCAAGTCCCAAT 1080  
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 DB 1081 ATCCAGGAGTTTGTCTGGCTGCAAGAAGATCTTTGGAGGCTGSCATTTCTGCCGAGAGC 1140  
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
 DB 1141 TTTGATGGGAGCCCAAGCCTCCAAACACTGCCCGCTCCAGCCAGAGCAGTCCAAGTGT 1200  
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420  
 DB 1201 GAGACTCTGGAAGAGATCACAGTTACCTATATACATCTCAGCATGGCCGAGACGCTGCCT 1260  
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 DB 1261 GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGAGCAATTCGACCAATGGCGCC 1320  
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 DB 1321 TACTCGCTGACCTGCAGGGCTGGGCATCAGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1380  
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480  
 DB 1381 CTGGGAGTGGACTGGCCCTCATCCACCATAACACCCACTCTGCTTCTGTCGACACGGTG 1440  
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 DB 1441 CCTGGGACCACTCTTTCCGAACCCGCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500  
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 DB 1501 GAGGACGAGTGTGTGGCGAGGGCTGGCTGCCACCACTGCTGCCCGCCGAGGCACTGC 1560  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 DB 1561 TGGGTTCAGGGCCCAACCCAGTGTGTCAACTGCACGAGTCTCTCTGGGGCCAGGAGTGC 1620  
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
 DB 1621 GTGGAGGAATCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680  
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 DB 1681 TTTGCCGTGCCACCTGAGTGTGAGGCCCAAGATGGCTCAGTAGCCTGTGTTTGGACCGGAG 1740  
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
 DB 1741 GCTGACCACTGTGTGGCTGTGCCCACTATAGGACCTCCCTTCTGCTGGCGCCGCTGC 1800  
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
 DB 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860  
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640

```
Db 1861 GCGCATGCCAGCCTTGCCCATCACTCACTGACCCACTCTCTGTGGACCTGGATGACAAG 1920
Qy 641 GlyCysProAlaGluInArgAlaSerProLeuThrSer----- 653
Db 1921 GGTCTCCCGCGGAGACAGAGACCGCCTCTGACGTCCATCATCTCTGCGGTGGTTGGC 1980
Qy 653 ----- 653
Db 1981 ATTCTGCTGCTGGTCTTGCGGTGCTCTTTGGGATCCCTCATCAAGCAGCGCAGCAG 2040
Qy 653 ----- 653
Db 2041 AAGATCCGGAAGTACACGATCGGAGACTGCTGCAGGAACGAGCTGGTGGAGCGCGTG 2100
Qy 653 ----- 653
Db 2101 ACACCTACGGGCGCATGCCCAACCAAGCGGCGAGATCGCGATCCTGAAAGACGCGAGCTG 2160
Qy 653 ----- 653
Db 2161 AGGAAGTGAAGTCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2220
Qy 653 ----- 653
Db 2221 CCTGATGGGAGAAATGTGAAATTCACGTGCCATCAAAGTGTGAGGGAAACACATCC 2280
Qy 653 ----- 653
Db 2281 CCCAAGCCAAACAAGAAATCTTAGCAAGACATACGTGATGGTGTGGCTCCCA 2340
Qy 653 ----- 653
Db 2341 TATGTCCTCCGCTTCTGGGCATCTGCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2400
Qy 653 ----- 653
Db 2401 ATGCCCTATGCTGCTCTTAGACATCTCGGGAAACCGCGGACGCTGGGCTCCAG 2460
Qy 653 ----- 653
Db 2461 GACCTGCTGAATGCTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATCTGGCG 2520
Qy 653 ----- 653
Db 2521 CTCGTACACAGGACTTGGCGCTCGGAACGTGCTGTCAAGAGTCCCAACCATGTCAA 2580
Qy 653 ----- 653
Db 2581 ATTACAGACTTCGGGCTGGCTCGGCTGGACATTTGACGACAGACAGATCATTGCAGAT 2640
Qy 653 ----- 653
Db 2641 GGGGCAAGGTGCCATCAAGTGGATGCGCTGGAGTCCATCTCCGCGCGGCTTACC 2700
Qy 653 ----- 653
Db 2701 CACCAGAGTGTGTGGAGTTATGTTGACTGTGTGGGAGCTGATGACTTTTGGGCGC 2760
Qy 653 ----- 653
Db 2761 AAACCTTACGATGGATCCAGCGCGGAGATCCCTGACCTGCTGGAAAGGGGAGCGG 2820
Qy 653 ----- 653
Db 2821 CTGCCCCAGCCCCCATCTGCACCATGATCTCTACATGATCATGGTCAAAATGTTGGATG 2880
Qy 653 ----- 653
Db 2881 ATTGACTCTGAATGTCGCGCAAGATTCGCGGAGTTGGTGTCTGAATTCCTCCCGATGCC 2940
Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
2941 AGGGACCCCGCGCTTGTGCTCATCCAGAATGAGGACTTGGGCCCGCAGCCAGTCCCTTG 3000
```

```
Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGCACCCTTCTACCGCTCACTGCTGGAGACGATGACATGGGGAGACTGGTGGATGCT 3060
Qy 685 GluGlyLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATCTGGTACCCAGCAGGGCTTCTTCTGTCAGACCTGCCCCGGCGCTGGG 3120
Qy 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
Db 3121 GGCATGGTCCACACAGCAGCCAGCTCATCTACAGAGTGGCGGTGGGACCTGACA 3180
Qy 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3181 CTAGGCTGGAGCCTCTGAAGAGGAGGCCCTCCACTGGCACCCTCCGAGGG 3240
Qy 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db 3241 GCTGGCTCCGATGATTTGATGGTGACCTGGGAATGGGGCAGCCAGGGCTGCAAGC 3300
Qy 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db 3301 CTCCCCACACATGACCCAGCCCTCTACAGCGGTACAGTGAGGAGCCACAGTACCCCTG 3360
Qy 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3361 CCTCTGAGACTGATGGCTAGCTTGCCCCCTGACCTGACGCCCCAGCCCTGAATATGTG 3420
Qy 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3421 AACGAGCAGATGTTGGCCCCAGCCCCCTTCGCCCGAGAGGGGCCCTCTGCTGCTGCC 3480
Qy 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3481 CGACTGCTGTGTGCACCTCTGGAAGGCCCAAGACTCTCTCCCCAGGGAAGAAATGGGGTC 3540
Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3541 GTCAAGACGTTTTTTCCTTTGGGGTGCCGTGGAGAACCCCGAGCTACTTGACACCCAG 3600
Qy 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3601 GGAGGAGTGCCTCCCTCAGCCCCACCTCTCTGCTTTCAGCCCGCAGCTTCGACAACTC 3660
Qy 885 TyrTyrTyrAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3661 TATTACTGGGACAGGAGCCACAGCGGGGGCTCCACCCAGCACCTTCAAGGGACA 3720
Qy 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3721 CCTACGCGAGAGAACCCAGACTACCTGGGTCTGGAGCTGCCAGTG 3765
```

## RESULT 2

```
US-08-356-786-1
; Sequence 1, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; STREET: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
```

; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA: US/08/356,786  
 ; APPLICATION NUMBER: US/08/356,786  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA: 07/831,967  
 ; APPLICATION NUMBER: 07/831,967  
 ; FILING DATE: 06-FEB-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Pitcher, Edmund R.  
 ; REGISTRATION NUMBER: 27,829  
 ; REFERENCE/DOCKET NUMBER: CRP-053  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 248-7000  
 ; TELEFAX: (617) 248-7100  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3768 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..3768  
 ; OTHER INFORMATION: /note= "product = "cerB-b2""  
 ; US-08-356-786-1

## Alignment Scores:

Pred. No.: 0 Length: 3768  
 Score: 4900.00 Matches: 919  
 Percent Similarity: 73.23% Conservative: 0  
 Best Local Similarity: 73.23% Mismatches: 0  
 Query Match: 96.49% Indels: 336  
 DB: 2 Gaps: 1

SEQ6 (1-919) x US-08-356-786-1 (1-3768)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProProGlyAla 20  
 DB 1 ATGGAGCTGGCGGCTGTGCGGCTGGGGCTCTCTCGCCCTCTTGGCCCGCGAGCC 60  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 DB 61 GCAGACCCCAAGTGTGCACCGGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 120  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
 DB 121 ACCCAGCTGGACATGCTCCGCCACCTCTACCCAGGGCTGCCAGGTGGTGCAGGGAAACCTG 180  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80  
 DB 181 GAATCAGCTTACCTGGCCCAACATGCGGCTCTCTCTCTCGAGGATATCCAGGAGGTG 240  
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 DB 241 CAGGGCTACGTGCTATCGCTCAACCAAGTAGGAGGCTCCCATCTGCAGAGGCTGCGG 300  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 DB 301 ATTGTGCGAGGCCACCCAGCTCTTTGAGGACAACATATGCTGCGCTGCTAGACAATGGA 360  
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 DB 361 GACCCGCTGAACAATAACACCCCTGTACAGGGGCTCCCGCAGGAGGCTGCGGGAGCTG 420  
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160  
 DB 421 CAGCTTCGAAGCTCAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCGAG 480

QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180  
 DB 481 CTCTGCTACCAGGACACGATTTTGTGAAGGACATCTCCACAAGAACAACACGCTGGCT 540  
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 DB 541 CTCACACTGATAGACACCAACCCCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAA 600  
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 DB 601 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGAGATTGTGAGAGCTGAGAGCTGAGCGCACTG 660  
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
 DB 661 GCCGGTGGCTGTGCCCGCTGCAGAGGGCCACTGCCACCTGACTGCTGCCATGAGCAGTGT 720  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 DB 721 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCCCTGCCCTTCCACTTCAACAC 780  
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 DB 781 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGTGTCACCTACACACACACAGCTTTGAG 840  
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 DB 841 TCCATGCCCAATCCCGAGGGCCGTATACATTCCGGCCCGCAGCTGTGTGACTGCTGTCCC 900  
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 DB 901 TACAACACTGCTTCTACGGACGTGGGATCCTGCAACCTCTGTCGCCCTGCACAACCAA 960  
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 DB 961 GAGGTGACAGCAGAGGATGAACACAGCGGTGTGAGAAGTGCAGAGCCCTGTGCCGA 1020  
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 DB 1021 GTGTGCTATGCTGGGCTGGAGCACTTGCAGAGGTGAGGCGAGTTACACAGTGCACAT 1080  
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 DB 1081 ATCCAGGAGTGTGCTGGCTGCAAGAAGATCTTTGGAGGCTGGCATTTCTGCCGGAGAGC 1140  
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400  
 DB 1141 TTTGATGGGAGCCACCCCTCCAACACTGCCCGCTCCAGCCAGGAGCTCCAGAGTGT 1200  
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420  
 DB 1201 GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCGGACAGCTTGCCT 1260  
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 DB 1261 GACCTCAGGCTCTTCCAGAACCTGCAAGTAATCCGGGGAGCAATCTGCACAATGGCGCC 1320  
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 DB 1321 TACTCCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGCTCCTCCTCAGGGAA 1380  
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480  
 DB 1381 CTGGGAGTGGAGCTGGCCCTCATCCACCATAACACCCACCTCTGCTTCTGTCACAGCGTG 1440  
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 DB 1441 CCTTGGAGCAGCTCTTTTCGGAACCCGACCAAGCTGTGCTCCACTTGCACACCGGCCA 1500  
 QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 DB 1501 GAGGACAGTGTGTGGCGGAGGCTTGGCTGCCACCACTGCTGCCCGGAGGAGGACTGC 1560  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540



D	b			TGGGGTCCAGGCGCCACCCAGGTGTCTCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC	1620
Q	y			VaIgluGlucysArgValLeuGlnGlyLeuProAraGluTyrrValAsnAlaAraHisCys	560
D	b			GTGGAGAAATGCCGAGTACTGCAGGGCTCCCAGGGAGTAGTGTAATGCCAGCACACTGT	1680
Q	y			LeuProCyHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
D	b			TTCCGTGGCACCCCTGAGTGTTCAGCCCCCAGAATGGCTCAGTGACCTGTTTTGGACCGAG	1740
Q	y			AlaAspGlnCysValAlaCysAlaHisTyrrLysAspProProPheCysValAlaAraCys	600
D	b			GCTGACAGTGTGTGGCTGTGCCACATATAAGGACCCTCCCTTCTGCGTGGCCCGGTGC	1800
Q	y			ProSerGlyValLysProAspLeuSerTyrrMetProIleTrpLysPheProAspGluGlu	620
D	b			CCCAGCGGTGAACCTGCACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
Q	y			GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640
D	b			GGCGCATGCCAGCCTTGGCCCCATCACTGCACCCACTCTGTGTGGACCTGGATGACAAG	1920
Q	y			GlyCysProAlaGluGlnArgAlaSerProLeuThrSer-----	653
D	b			GGTGTCCCCCGGAGCAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGC	1980
Q	y			-----	653
D	b			ATTCTGCTGGTGGTCTTGGGGTGGTCTTTGGGATCCTCATCAACGAGCGCAGCAG	2040
Q	y			-----	653
D	b			AAGATCCGGAAGTACACGATGCGGAGACTGTCTGCAGGAACCGGAGCTGTGTGGAGCCGCTG	2100
Q	y			-----	653
D	b			ACACTAGCGGAGCGATGCCAACCGCGCGCAGATGCGGATCCTGAAAGAGACGGAGCTG	2160
Q	y			-----	653
D	b			AGGAAGGTGAAGTGTGGATCTGGCGCTTTTGGCACAGCTCTACAGGGGATCTGGATC	2220
Q	y			-----	653
D	b			CCTGATGGGAGAAATGTGAAATTTCCAGTGGCCATCAAAGTTTGAGGGAAAAACACATCC	2280
Q	y			-----	653
D	b			CCAAAAGCCAAGAAATCTTAGACGAAGCATACGTGATGGTGTGGTGGCTCCCCCA	2340
Q	y			-----	653
D	b			TATGTCTCCGGCTTCTGGGCATCTGCCCTGACATCCAGGTGCAGCTGGTGACACAGCTT	2400
Q	y			-----	653
D	b			ATGCCCTATGGTGCCTCTTAGACCATGTCCGGGAAACCGGGACGCCCTGGGCTCCAG	2460
Q	y			-----	653
D	b			GACCTGCTGAAGTGTGTATGACGATTTGCCAAGGGGATGAGTACCTGGAGGATGTGCGG	2520
Q	y			-----	653
D	b			CTCGTACACAGGACTTGGCGCTCGGAACGTGTGGTCAAGAGTCCCAACCATGTCAAA	2580
Q	y			-----	653
D	b			ATTACAGACTTCGGGCTGCTCGGCTGTGGNACATTGACGACAGACAGATACCATTGCAGAT	2640
Q	y			-----	653

```

RESULT 3
US-09-048-804-1
; Sequence 1, Application US/09048804
; Patent No. 5968748
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
; TITLE OF INVENTION: HUMAN HER-2 EXPRESSION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748ris LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/048,804
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2913
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4473 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; ANTI-SENSE: No
US-09-048-804-1

Alignment Scores:
Pred. No.: 0 Length: 4473
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 2 Gaps: 1

SEQ6 (1-919) x US-09-048-804-1 (1-4473)
QY 1 MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 175 ATGGAGCTGGCGGCTTGCGCGCTGGGGGCTCCCTCGCCCTCTTGCCCGCCGGAGCC 234
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCGAGACCCAGTGTGACCGGCACAGACATGAGCTGGGCTCCCTGCCAGTCCCGAG 294
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 295 ACCACCTGGACATGCTCGCCACCTCTACCAAGGCTGCCAGGTGGTGGAGGAAACCTG 354
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 355 GAACACACCTACTGCTCCCAACATGCCAGCCTGCTCTCTCCAGGATATCCAGAGGTG 414
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGGCTAGCTGCTCATGCTCACACCAAGTGAAGGAGGCTCCCACTGAGAGGCTGCGG 474
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120

```

```

Db 475 ATTGTGCGAGGCACCCAGCTCTTTGAGGACAACACTATGCCCTGCCGTGTACAGCAATGA 534
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGCTGACAATATACCACTCCCTGTACAGGGGCTCCCAAGAGGCTGCGGGAGCTG 594
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 595 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGGGGCTTGTATCAGGGGAACCCCGAG 654
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 655 CTCTGCTACCAAGGACACGATTTTGTGAAGGACATCTTCCACAAGAACCAACAGCTGG 714
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 715 CTCACACTGATAGACCAACCGCTCTCGGGCTGCCACCTGTCTCCGATGTGTAAG 774
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 775 GGCTCCCGCTGCTGGGAGAGAGATTCTGAGGATTTGTACAGGCTGACGGGCACTGCTGT 834
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCys 240
Db 835 GCGGCTGGCTGTGCCCGCTGCAAGGGGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 895 GCTCCCGCTGCACGGGCCCCAAGCACTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 955 AGTGCATCTGTGAGCTGCACCTGCCAGCCCTGTGAGTACCTACAAACACAGACAGCTT 1014
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 1015 TCCATGCCCAATCCCGAGGCGGTATACATTCCGGCGCAGCTGTGTGCTGCTGCTGCTG 1074
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1075 TACAACTACCTTTCTACGGACGTGGGATCCTGCACCCCTGCTGCTGCTGCTGCTGCTG 1134
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1135 GAGGTGACAGCAGAGAGATGGAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCGA 1194
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1195 GTGTGCTATGTGTGGCATGGAGCACTTGGCAGAGGTGAGGGCAGTTACCAAGTGCCT 1254
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1255 ATCCAGGAGTTGTGTGGCTGCAAGAGATCTTTGGAGCTGGCATTTCTGCCGGAGAGC 1314
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1315 TTTGATGGGAGCCAGCCCTCCAACTGCCCGCTCCAGCCAGCAGCAGCTCCAGTGT 1374
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1375 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGGCCGAGCAGCTGCCT 1434
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1435 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCCGGGGAGCAATTTCTGCACAAATGG 1494
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1495 TACTCGCTGACCTGCAGGGCTGGGCATCAGCTGCTGGTGGGGCTGGCTGCTGCTGCTG 1554
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480

```

Db 1555 CTGGGAGTGAGCTGGGCGCTCATCCACATCAACACCCACCTCTCTCGTGCACACGGTG 1614

Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
|||||

Db 1615 CCCTGGGACAGCTCTTTCGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1674

Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
|||||

Db 1675 GAGGACGAGTGTGGGCGAGGCGCTGGCGCTGCCACAGCTGTGGCGCGAGGCGACTGC 1734

Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
|||||

Db 1735 TGGGGTCCAGGGCCCGCCAGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1794

Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
|||||

Db 1795 GTGGAGGAAATGCCAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1854

Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
|||||

Db 1855 TTGCGGTGCCACCTGAGTGTGACGCCCGCAGAAATGGCTCAGTGACCTGTTTTGGACCGGAG 1914

Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrIlyAspProPheCysValAlaArgCys 600  
|||||

Db 1915 GCTGACCAAGTGTGGGCTGTGCCCACTATAAGACCTCCCTTCGCTGGCGCGCTGC 1974

Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
|||||

Db 1975 CCCAGCGGTGGAACCTGACCTCTCTACATGCCATCGCCATCTGGAAGTTTCCAGATGAGGAG 2034

Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640  
|||||

Db 2035 GGCGCATGCCAGCCTTGCCCATCACTGCACCCACCTCCTGTGTGGACCTGGATGACAAG 2094

Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer ----- 653

Db 2095 GGCTGCCCGCGGACGAGAGAGCCGCTCTGACGTCCATCATCTCTCGCGGTGGTTGGC 2154

Qy 653 ----- 653

Db 2155 ATTCGTGCTGGTGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGCACAG 2214

Qy 653 ----- 653

Db 2215 AAGATCCGGAAGTACACGATCGGAGACTGTGCGAAGAACGGAGCTGGTGGACCGCTG 2274

Qy 653 ----- 653

Db 2275 ACACCTAGCGGAGCGATGCCCCAACCCAGGCGAGATCGGATCCTGTAAAGAGACGGAGCTG 2334

Qy 653 ----- 653

Db 2335 AGGAAGGTGAAGGTGGTCTGGGCGCTTTTGGCACAGTCTACAGGGCATCTGGATC 2394

Qy 653 ----- 653

Db 2395 CCTGATGGGAGAGATGTGAAATTCACAGTGCCCATCAAGTGTGTGGGAAACACATCC 2454

Qy 653 ----- 653

Db 2455 CCCAAAGCCAAAGAAATCTTAGACGAAGCATACGTGATGCGTGTGGTGTGCCCTCCCA 2514

Qy 653 ----- 653

Db 2515 TATGTCTCCGCGCTTCTGGGCATCTGCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2574

Qy 653 ----- 653

Db 2575 ATGCCCTATGGCTGCCCTTTAGACCATGTCCGGGAAACCGCGACGCTGGGTGCCAG 2634

Qy 653 ----- 653

Db 2635 GACCTGCTGAAGTGTGTATGCAGATTTGCCAAGGGGATGAGTACCTGGAGAGATGTGCGG 2694

Qy 653 ----- 653

Db 2695 CTCTACACAGGGACTTGGCGCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAAA 2754

Qy 653 ----- 653

Db 2755 ATTACAGACTTTCGGGCTGGCTCGGCTGTGGACATTTGACGAGACAGAGTACCATGCAGAT 2814

Qy 653 ----- 653

Db 2815 GGGGCAAGTGCCCATCAAGTGGCTGGAGTCCATTTCTCCGCGCGGTTTCACC 2874

Qy 653 ----- 653

Db 2875 CACCAGAGTGTGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC 2934

Qy 653 ----- 653

Db 2935 AAACCTTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGGAGCGG 2994

Qy 653 ----- 653

Db 2995 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATPCATGGTCAAAATGTTGGATG 3054

Qy 653 ----- 653

Db 3055 ATTGACTCTGAATCTCGGCCAAGATTCGGGAGGTTGGTGTCTGAATTTCTCCCGCATGGCC 3114

Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664  
|||||

Db 3115 AGGACCCCCAGCGCTTTTGGTCAATCAGAAATCAGGACTTGGGCCCGCCAGCTCCCTTG 3174

Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684  
|||||

Db 3175 GACAGCACTTCTACCCCTCCTGCTGGAGACATGACATGGGGGACCTGGTGGATGCT 3234

Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704  
|||||

Db 3235 GAGGAGTATCTGGTACCCCGAGCGGCTTCTTGTCCAGACCTGCCCGGGCGCTGGG 3294

Qy 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724  
|||||

Db 3295 GGCATGTCCACACAGCGCACCGCAGCTCATCTACAGGAGTGGCGGTGGGACCTGACA 3354

Qy 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744  
|||||

Db 3355 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCGCCCAAGTCTCCACTGGCACCTCCGAAGG 3414

Qy 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnIns 764  
|||||

Db 3415 GCTGGCTCCGATGTATTTGATGGTACCTGGGAATGGGGCAGCGCAAGGGCTCCAAAGC 3474

Qy 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
|||||

Db 3475 CTCCCCACATACCCCGCCCTCTACAGCGGTACAGTGGAGGACCCACAGTACCCTG 3534

Qy 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
|||||

Db 3535 CCCTCTGAGACTGATGGCTACGTTGGCCCCCTGACCTGCAGCCCCCGCCAGCTGAATATGTG 3594

Qy 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824  
|||||

Db 3595 AACCGCCAGATGTTCCGCCCCAGCCCCCTTCGCCCCGAGAGGGCCCTCTGCTGTGTC 3654

Qy 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
|||||

Db 3655 CGACCTGCTGGTCCACTCTGGAAGGCGCCCAAGACTCTCTCCCGAGGAAAGATGGGGTC 3714

Qy 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
|||||

Db 3715 GTCAAGACGCTTTTTCCTTTGGGGTGGCTGTGAGAAACCCCGAGTACTTTGACACCCCG 3774



Db 1675 GAGGACGAGTGTGGGGCAGGGGCTGGCCCTGCCACACAGCTGTGGCCCGGAGGCACTGC 1734  
 Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 Db 1735 TGGGGTCCAGGGCCACCCAGTGTCAACTGACAGCCAGTTCTTGGGGCCAGAGTGC 1794  
 Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaAraHisCys 560  
 Db 1795 GTGGAGGAATCCCGAGTACTCAGGGGCTCCAGGGAGTATGTGAATGCCAGGCACTGT 1854  
 Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 Db 1855 TTGGCGTGCCACCTGTAGTGTGACGCCCGAGAAATGGCTGAGTACCTGTTTGGACCGGAG 1914  
 Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
 Db 1915 GCTGACCAGTGTGGCCTGTGCCCACTATPAAAGACCTCCCTTCTGGTGGCCCGCTGC 1974  
 Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProLeuTyrLysPheProAspGluGlu 620  
 Db 1975 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGTAAGTGTCCAGATGAGGAG 2034  
 Qy 621 GlyAlaCysGlnProCysProLeuAsnCysThrHisSerCysValAspLeuAspAspLys 640  
 Db 2035 GCGCATGCCAGCCTTGCCCATCACTGACACCCACCTCTGTGTGGACCTGGATGACAAG 2094  
 Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
 Db 2095 GGCTGCCCGCGAGAGAGAGCCGCTCTGACGTCCATCTCTGCGGTGGTTGGC 2154  
 Qy 653 2155 ATCTGTGTCGTGGTCTTGGGGTGTCTTGGGATCCTCATCAAGCAGCGCAGCAG 2214  
 Qy 653 2215 AAGATCCGGAGTACACGATGCGGAGACTGCTGAGGAAACGGAGCTGGTGGAGCCGCTG 2274  
 Qy 653 2275 ACACCTAGCGAGCGATGCCCAACCCAGCGCAGATCGGATCCTGAAAGACAGCGAGCTG 2334  
 Qy 653 2335 AGGAAGTGAAGGTCTTGGATCTGGCCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2394  
 Qy 653 2395 CCTGATGGGGAGAAATGTGAAATTTCCAGTGGCCATCAAAAGTGTGAGGGAACACATCC 2454  
 Qy 653 2455 CCCAAGCCACAAGAATCTTAGCAAGCATACGTGATGGCTGGTGGTGGGCTCCCA 2514  
 Qy 653 2515 TATGTCCTCCGCTTCTGGGCATCTGCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2574  
 Qy 653 2575 ATGCCCTATGGCTGCCCTTAGACCATGTCCGGGAAACCGCGGACGCCCTGGGCTCCAG 2634  
 Qy 653 2635 GACCTGTGAACCTGTATGCAGATTTCCNAGGGGATGAGCTACCTGGAGGATGTCGG 2694  
 Qy 653 2695 CTCGTACACAGGACTTGGCGCTCGGAACGTGCTGCTCAAGAGTCCCAACCATGTCAAA 2754  
 Qy 653 2755 ATTACAGACTTCGGGCTGGCTGCTGACATTTGACGACATGACGAGACAGAGTACCATGCAGAT 2814

Qy 653 2815 GGGGCAAGGTGCCCATCAAGTGGATGCCGCTGGAGTCCATCTCCCGCGGGGTTTACC 2874  
 Db 653 2875 CACCAGAGTGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2934  
 Qy 653 2935 AAACCTTACGATGGATGCCAGCCCGGAGATCCCTGACCTGCTGGAAAAAGGGAGCGG 2994  
 Db 653 2995 CTGCCCCAGCCCCCTCTGCACCATTTGATGCTTACATGATCATGGTCAAAATGTTGGATG 3054  
 Qy 653 3055 ATTGACTCTGAATGTCGGCCCAAGATTCCGGGAGTTGGTGTCTGAATTTCTCCCGCATGGCC 3114  
 Qy 654 3115 AGGAAACCCAGCCCTTTGTGGTCAATCAGAAATGAGGACTTGGGCCCCAGCCAGTCCCTTG 3174  
 Db 664 3175 GACAGACCTTCTACCCGCTCACTGCTGGAGGACCATGACATGGGGACCTGGTGGATGCT 3234  
 Qy 684 3235 GAGGAGTATCTGGTACCCAGCAGGGCTTCTTCTGTCAGACCTGCCCCGGCGCTGGG 3294  
 Db 704 705 GlyMetValHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724  
 Db 724 3295 GGCATGTCTCACCACAGCCACCGAGCTCACTACAGGAGTGGCGGTGGGAGCTTGACA 3354  
 Qy 744 3355 CTAGGGCTGGAGCCCTCTGAAGAGAGGCCCCCAGGCTCTCCACTGGCACCTCCGAAGGG 3414  
 Db 764 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764  
 Db 3415 GCTGGCTCCGATGATTTGATGTGACCTGGGAATGGGGCAGCACAAGGGCTGCAAGC 3474  
 Qy 784 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
 Db 3534 3475 CTCCCCACATGACCCCGCCCTCTACGCGGTACAGTGGAGGACCCACAGTACCCCTG 3534  
 Qy 804 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
 Db 3594 3535 CCTCTGAGACTGATGCTAGCTGCCCCCTGACCTGACGCCCCCAGCTGGAATATGTG 3594  
 Qy 824 805 AsnGlnProAspValArgProGlnProSerProArgGluGlyProLeuProAlaAla 824  
 Db 3654 3595 AACCCAGCAGATGTTGGCCCCCAGCCCCCTTCGCCCGCAGAGGGGCCCTCTGCCCTGCTGC 3654  
 Qy 844 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
 Db 3714 3655 CGACCTGCTGGTGCATCTGGAAGGCCCCAAGACTCTCTCCCCAGGGAAGATGGGGTC 3714  
 Qy 864 845 ValLysAspValPheAlaPheGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
 Db 3774 3715 GTCAAAACGCTTTTGGCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGCACACCCAG 3774  
 Qy 884 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884  
 Db 3834 3775 GGAGGAGTGCCTCAGCCCCACCTCTCTCTGCTTCCAGCCCCAGGCTTTCGACAACTC 3834  
 Qy 904 885 TyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
 Db 3894 3835 TATTACTGGGACGAGGACCCAGGAGGGGGCTCCACCCAGCACCTTCAAGAGGACA 3894



QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480  
 DB 1531 CTGGCAGTGGAGCTGGCCCTCATCCACCATACACCCACCTCTGCTGTCACACGGTG 1590  
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 DB 1591 CCTCGGGACCAAGCTCTTTCCGAAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCA 1650  
 QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 DB 1651 GAGGACGAGTGTGGGGAGGGCCCTGGCTGCCACAGCTGTGCGCCGAGGGCACTGC 1710  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 DB 1711 TGGGTCCAGGGCCACCCAGTGTGTCAACTGCAGCCAGTCTCTTCGGGGCCAGGAGTGC 1770  
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
 DB 1771 GTGGAGGAAATGCCGAGTACTGCGAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACGTGT 1830  
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 DB 1831 TTGCGGTGCCACCCCTGAGTGTGAGCCCGCAGAAATGCTCAGTGACCTGTTTGGACCGGAG 1890  
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrIlysAspProProPheCysValAlaArgCys 600  
 DB 1891 GCTGACCAAGTGTGGCTGTGCCCTATATAAGGACCCCTCCCTTCTGCGTGGCCGCTGC 1950  
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
 DB 1951 CCCAGCGGTGAAACCTTGACCTCTCCACATGCCCATCTGGAAGTTCCAGATGAGGAG 2010  
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640  
 DB 2011 GCGCATGCCAGCCTTGCCCATCACTGCACCCACTCTGTGTGGACCTGGATGACAAAG 2070  
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
 DB 2071 GGCTGCCCGCGGACAGAGAGCCAGCCCTCTGAGTCCATCTCTGCGGTGTTGGC 2130  
 QY 653 653 653  
 DB 2131 ATTCTGCTGGTGGTCTTGGGGTGGTCTTTGGGATCTCATCAACGCGAGCGAGCAG 2190  
 QY 653 653 653  
 DB 2191 AAGATCCGGAAGTACACGATCGGAGACTGCTGCGAGAAACGGAGCTGGTGGAGCCGCTG 2250  
 QY 653 653 653  
 DB 2251 ACACCTAGCGGAGCGATCCCAACACGCGCAGATGCGGATCTCTGAAAGAGACGGAGCTG 2310  
 QY 653 653 653  
 DB 2311 AGGAAGGTGAAGGTCTTGGATCTGGCCCTTTTGGCACAGCTACAGGGGATCTGGATC 2370  
 QY 653 653 653  
 DB 2371 CCTGATGGGAGAAATGTGAAATTCAGTGGCCATCAAGTGTTCAGGGAAACACATCC 2430  
 QY 653 653 653  
 DB 2431 CCCAAAGCCAAAGAAATCTTAGCAAGCATACGTGATGGTGTGGTGTGGCTCCCCA 2490  
 QY 653 653 653  
 DB 2491 TATGTCTCCGCCCTTCTGGGCATCTGCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2550  
 QY 653 653 653  
 DB 2551 ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCCGCGAGCCTGGGCTCCAG 2610

QY 653 653 653  
 DB 2611 GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTTGAGGATGTGCGG 2670  
 QY 653 653 653  
 DB 2671 CTGCTACACAGGGACTTGGCCCTCGGAACGTCGCTCAAGAGTCCCAACCATGTCAA 2730  
 QY 653 653 653  
 DB 2731 ATTACAGACTTCGGGCTGGCTGGCTGGACATTTGACGACAGAGTACCATCGAGAT 2790  
 QY 653 653 653  
 DB 2791 GGGGCAAGGTGCCCATCAAGTGGATGGCTGGAGTCCATTCTCCGCGGGGGTTTACC 2850  
 QY 653 653 653  
 DB 2851 CACCAGAGTATGTGTGGAGTTATGGTGTGAGTGTGTGGGAGCTGATGACTTTTGGGGCC 2910  
 QY 653 653 653  
 DB 2911 AAACCTTACGATGGGATCCCGCCGCGGAGATCCCTGACCTGCTGGAAGGGGGAGCGG 2970  
 QY 653 653 653  
 DB 2971 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGGTCAAAATGTTGGATG 3030  
 QY 653 653 653  
 DB 3031 ATTGACTCTGAATGTGCGCCAAGATTCCGGGAGTTGGTGTCTGAATTTCCCGCATGCCC 3090  
 QY 654 654 654  
 DB 3091 AGGGACCCCCAGCGCTTTGTGTCATCCAGATGAGGACTTGGGCCCCAGCCAGTCCCTTG 3150  
 QY 665 AspSerThrPheTyrArgSerLeuLeuAspAspMetGlyAspLeuValAspAla 684  
 DB 3151 GACAGCACTTCTACCGCTCACTGTGTGGAGACGATGATGTTGGGAGCTGTGTGATGCT 3210  
 QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704  
 DB 3211 GAGGAGTATCTGGTACCCCGACGAGGCTTCTTCTGTCAGACCTTCCCGCGGCGCTGGG 3270  
 QY 705 GlyMetValHisHisArgSerSerThrArgSerGlyGlyAspLeuThr 724  
 DB 3271 GGCATGGTCCACACAGCAGCCAGCTCATCTACAGAGTGGCGGTGGGACCTGACA 3330  
 QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744  
 DB 3331 CTAGGGCTGGAGCCCTCTGAAGAGAGAGGCCCGCCAGGTCTCCACTGGCACCTTCCGAAGG 3390  
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764  
 DB 3391 GCTGCTCCGAGTATTTGATGGTGACCTGGGAATGGGGGAGCCCAAGGGCTGCAAGAGC 3450  
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
 DB 3451 CTCCCCACACATGACCCCGAGCCCTCTACAGCGGTACAGTAGGAGCCACACAGTACCCCTG 3510  
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
 DB 3511 CCTCTGAGACTGATGGCTAGCTTGGCCCTTGGCTGACCTGACGCCCCCAGCCCTGAATATGTG 3570  
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824  
 DB 3571 AACCGCCAGATGTTCGGCCCGAGCCCTTTCGCCCGAGAGGGGCCCTTCTCCCTGCTGCC 3630  
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
 DB 3631 CGACCTGTGTGCCACTCTGGAAGGGCCAGAGACTCTCTCCCGAGGAAAGTGGGGTGC 3690  
 QY 845 ValLysAspValPheAlaPheGlyAlaValGluAsnProGluTyrLeuThrProGln 864

```

|||||
3691 GTCAAGACGCTTTTGGCTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCAG 3750
QY
865 GYGlyAlaalaProGlnProHisProProAlaPheSerProAlaPheAsnLeu 884
Db
3751 GGAGGAGCTGCCCTCAGCCCCACCCTCTCTCCCTTCAGCCCGCCCTTCGACAACCTC 3810
QY
885 TTTTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db
3811 TATTACTGGACACAGACCACACAGAGCGGGGGCTCCACCCAGACCTTCAAAAGGACA 3870
QY
905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db
3871 CCTAGCGACAGAAACACAGAGTACCTGGGTCTGGAGCTGCCAGTG 3915

RESULT 6
US-08-645-865-9
; Sequence 9, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPlastic DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-645-865-9

Alignment Scores:
Pred. No.: 0 Length: 4530
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservative: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 1 Gaps: 1

SEQ6 (1-919) x US-08-645-865-9 (1-4530)
QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProProGlyAla 20
Db 151 ATGGAGCTGGCGGCTTGTGGCGCTGGGGCTCCTCTCGCCCTCTTGGCCCCCGGAGCC 210
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 GCGAGCACCCAGTGTGACCCGCGACAGACATGAAGCTGGGCTCCTCGCCAGTCCCGAG 270

```

```

QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCCACTGGACATGTCTCCGCCACCTCTACCAAGGGTGCAGAGTGGTGCAGGAGAACTG 330
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 331 GAACTCACCTTACCTGCCCAATGCCACCAATGCCAGCTCTCTCTCTGCTGAGGATATCCAGGAGGTG 390
QY 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGTACGTGCTCATCGCTCAACCAAGTAGGAGGCTGCCACTGCAGAGGCTGCGG 450
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGCGAGGACCCAGCTCTTTGAGGACACTATGCCCTGCCCTGCTGAGACATGGA 510
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGCTGAACAATACCAACCCCTGTACACAGGGGCTTCCCAAGAGGCTGCGGAGCTG 570
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuLeuGlnArgAsnProGln 160
Db 571 CAGCTTCGAAGCCTCAGAGATCTTGAAGAGGAGGCTCTTGATCCAGCGGAACCCCGAG 630
QY 161 LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeuAla 180
Db 631 CTCTGCTACGAGCACAGATTTTGTGGAAGGACATCTTCCCAAGAACACACAGCTGGCT 690
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACACCAACCGCTCTCGGGGCTGCCACCCCTGTCTCCGATGTGTAAG 750
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCCGCTGCTGGGAGAGATTTCTGAGGATTTGTCAGAGCTGACGCGCATCTGTGT 810
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 811 GCGGGTGGCTGTGCCCGCTGCAAGGGCCACTGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 870
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 871 GCTGCCGGTGCACGGGCGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 931 AGTGGCATCTGTGAGCTGCACTGCCCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 990
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 991 TCCATGCCCAATCCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGCTGCTGCTGCTGCT 1050
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1051 TACAACCTACCTTTCTACGGACGTGGATCTCTGACCCCTGCTGCTGCTGCTGCTGCTGCTGCT 1110
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGlyLysSerLysProCysAlaArg 340
Db 1111 GAGGTGACAGCAGAGGATGAACACAGCGGTGTGAGAAGTGCAGAGGCTGTGCGCCGA 1170
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1171 GTGTGCTATGCTGCGCATGGAGCACTTGCAGAGGTGAGGCGAGTACCAGTGCCAAT 1230
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1231 ATCCAGAGTGTCTGGCTGCAAGAAGATCTTTTGGAGCGCTGGCATTTCTGCGCGGAGAG 1290
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1291 TTTGATGGGAGCCAGCCCTCCAACACTGCCCCGCTCCAGCCAGACAGCTCCCAAGTGTTT 1350

```



Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1351		
		GAGACTCTGGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCGGACAGCGCTGCCT	1410
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1411		
		GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGACGAATCTGCACAATGGCGGCC	1470
Qy	441	TyrSerLeuThrLeuGlnGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1471		
		TACTCGCTAGCCCTGCAAGGCGCTGGGCATCAGCTGGCTGGGCGTGCCTCACTCAGGGA	1530
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1531		
		CTGGCAGTGGACTGGCCCTCATCCACATAAACCCACCCTCTCTCTGTCGACACGCGTG	1590
Qy	481	ProTrpAspGlnLeuPheArgAspProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1591		
		CCCTGGNACCAGCTCTTTCCGAACCCGACCAAGCTCTGCTCCACACTGCCCAACCGGCCA	1650
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1651		
		GAGGACGAGTGTGTGGCGAGGGCTGGCCTGCCACCAGCTGTGCGCCCGAGGGGCACTGC	1710
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1711		
		TGGGTCGAGGGCCACCCAGTGTCAACTGCAGCCAGTTCCCTTCGGGGCCAGGAGTGC	1770
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1771		
		GTGGAGGAATGCCGAGTACTGACGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACGTG	1830
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1831		
		TTGCGTGCACCCCTGAGTGTGACGCCCCAGNAATGGCTCAGTGACCTGTTTGGACCGGAG	1890
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1891		
		GCTGACCAGTGTGTGGCTGTGTGCCACTATAAGGACCCCTCCCTTCTGCTGGCGCCGCTGC	1950
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1951		
		CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGTGAAGTTTCAGATGAGGAG	2010
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2011		
		GGCGCATGCCAGCTTGCCCATCACTGCACCCACCTCCTGTGTGGACCTGGATGACAG	2070
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2071		
		GGCTGCCCGCGGACGAGAGAGCCAGCCCTCTGACGTCCTCATCGTCTCTGCGGTGTTGGC	2130
Qy	653	-----	653
Db	2131	ATTCTGCTGTCGTGGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAG	2190
Qy	653	-----	653
Db	2191	AAGATCCGGAAGTACAGATCGCGAGACTGCTGCAGGAACCGGAGCTGTGGAGCCCGCTG	2250
Qy	653	-----	653
Db	2251	ACACCTAGCGGAGCGATGCCCAACCGCGCAGATCGCGATCTCTGAAAGAGACGGAGCTG	2310
Qy	653	-----	653
Db	2311	AGGAAGGTGAAGTGTGTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2370
Qy	653	-----	653
Db	2371	CCTGATGGGGAAGATGTGAATAATTCAGTGGCCATCAAGATGTGAGGGGAAAAACATATCC	2430
Qy	653	-----	653

Db 3511 CCTCTGAGACTGATGGCTACGTTGCCCTTACCTGACGAGCCCGAGCTGAATATGTG 3570  
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824  
 Db 3571 AACCCACAGATGTTGGCCCGCCAGCCCTTGGCCCGAGAGGGCCCTCTGCTGTGCTGCC 3630  
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
 Db 3631 CGACCTGCTGGTCCCACTCTGGAAGGGCCCAAGACTCTCTCCCGAGGAAGTGGGTC 3690  
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
 Db 3691 GTCAAGAGCGTTTTCCTTGGGGTGGCTGGAGAACCCCGAGTACTTGACACCCAG 3750  
 QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884  
 Db 3751 GGAGGAGCTGCCCTCAGCCCACTCTCTCTGCTTCCAGCCAGCTTCGACAACCTC 3810  
 QY 885 TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
 Db 3811 TATTACTGGGACCAAGCACCAGAGCGGGGGCTCCACCCAGCACCTTCAAAGGGACA 3870  
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
 Db 3871 CCTACGGCAGAGAACCAGAGTACCTGGGTCTGGAGCTGCCAGTG 3915

RESULT 7

US-09-167-322-4  
 : Sequence 4, Application US/09167322  
 : Patent No. 6365151  
 : GENERAL INFORMATION:  
 : APPLICANT: Allegheny University of the Health  
 : Sciences, Halpern, Michael S.  
 : England, James M.  
 : TITLE OF INVENTION: CANCER VACCINE  
 : NUMBER OF SEQUENCES: 14  
 : CORRESPONDENCE ADDRESSES:  
 : ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.  
 : STREET: Suite 1800, Two Penn Center Plaza  
 : CITY: Philadelphia  
 : STATE: PA  
 : COUNTRY: USA  
 : ZIP: 19102

COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/09/167,322  
 : FILING DATE: 07-Oct-1998  
 : CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: PCT/US97/00582  
 : FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:  
 : NAME: Monaco, Daniel A.  
 : REGISTRATION NUMBER: 30,480

REFERENCE/DOCKET NUMBER: 7933-33 PC  
 : TELEPHONE: (215) 568-8383  
 : TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 4:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 4530 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
 : US-09-167-322-4

Alignment Scores: 0 Length: 4530  
 Pred. No.:

Score: 4892.00 Matches: 918  
 Percent Similarity: 73.15% Conservative: 0  
 Best Local Similarity: 73.15% Mismatches: 1  
 Query Match: 96.34% Indels: 336  
 DB: 4 Gaps: 1

SEQ6 (1-919) x US-09-167-322-4 (1-4530)

QY	1	MetGluLeuAlaAlaLeuCysArgTTPGlyLeuLeuAlaLeuProProGlyAla	20
Db	151	ATGAGCTGGCGCCCTGTGCGCTGGGGCTCCTCTCGCCCTCTTGCCCGGAGCC	210
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	211	GGGAGCACCAAGTGTGCACCGGCACAGACATGAAGTGGGGTCCCTGCCAGTCCGAG	270
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60
Db	271	ACCCAGCTGGACATGCTCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGGAACCTG	330
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	331	GAACCTACCTACCTGCCACCAATGCCAGCTGCTCTCTCGAGGATATCCAGGAGGTG	390
QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	391	CAGGGCTACGTGCTCATGCTCAACAACCAAGTGAAGGAGGGTCTTGATCCAGGGAA	450
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	451	ATTGTGCGAGGCAACCAGCTCTTTGAGGACAACTATGCCCTGGCCCTGTAGACAATGA	510
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Db	511	GACCCGCTGAACAAATACCAACCCCTGTCAAGGGCCCTCCCGAGGAGGCTGGCGGAG	570
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160
Db	571	CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGTCTTGATCCAGGGAAACCCCA	630
QY	161	LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAla	180
Db	631	CTCTGTACAGGACACGATTTGTGGAAGGACATCTCCACAAGAACACACAGCTGGCT	690
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	691	CTCACACTATAGACACCAACCCCTCTCGGGCTGCCACCCCTCTGCTCTCGCATGTAG	750
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	751	GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTACAGAGCTGACGCGCACTGT	810
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGluGlnCys	240
Db	811	GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGCTGCTGCCATGACAGTGT	870
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	871	GCTGCCGGTGCACGGGGCCCAAGCACTCTGACTGCTGGCTGCCCTGCCCTCCACTT	930
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	931	AGTGGCATCTGTGAGCTGCACCTGCCCGAGCCCTGGTCACTACACACACACACAG	990
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	991	TCCATGCCCAATCCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGTGCTGTCC	1050
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	1051	TACAACTACCTTTCTACGGACGTGGGATCTTGACCCCTCTGTCACCCCTCTGCA	1110
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340





Db	1231	ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTTGGCATTTCTGCCGGAGAGC	1290
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1291	TTTTGATGGGACCCAGAGCTCCACACGTCCGCCCGCTCCAGCAGAGCAGCTCCAAGTGT	1350
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1351	GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGAGCAGCTGCCT	1410
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1411	GACCTCAGCGCTTTCCAGAACCTGCAAGTAATCCGGGACGAATCTCGCAATGGCGCC	1470
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1471	TACTCGCTGACCTTGCAGGGCTGGGCATCAGCTGGCTGGGCTGGCCCTCACTGAGGGAA	1530
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1531	CTGGCAGTGAGCTGGCCCTCATCCACCATAAACACCCACCTCTGCTTCGTGCACACGGTG	1590
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1591	CCCTGGGACCAAGCTCTTTCCGAACCCGACCAAGCTCTGTCTCCACACTGCCAACCGGCCA	1650
Qy	501	GluAspGluCysValGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1651	GAGGACGAGTGTGGCGGAGGGCTGGCCCTGCCACCAAGCTGTGGCCCGGAGGCACTGC	1710
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1711	TGGGGTCCAGGGCCACCCAGTGTCACCTGACGACCCAGTTCCTTCGGGGCCAGGAGTGC	1770
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1771	GTGGAGGAATGCCAGTACTGTCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1830
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1831	TTGCCGTGCCACCTGAGTGTACGCCCCAGAAATGGCTCAGTACCTGTTTGGACCGGAG	1890
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1891	GCTCACCAGTGTGGCGCTGTGCCCACTATAAGGACCCCTCCCTTCGCGTGGCCCGCTGC	1950
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1951	CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	2010
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2011	GGCGCATGCCAGCTTGCCCCATCACTGACCCCACTCCCTGTGTGGACCTGGATGACAAG	2070
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2071	GGCTGCCCCCGCAGCAGAGAGCCAGCCCTCTGAGGTCCTCTGCGGTGGTTGGC	2130
Qy	653	-----	653
Db	2131	ATTCTGTGTCGTGGTCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGCAGCAG	2190
Qy	653	-----	653
Db	2191	AAGATCCGGAAGTACAGATGGGAGACTGCTGCAGGAAACAGGAGTGGTGGAGCCCGCTG	2250
Qy	653	-----	653
Db	2251	ACACCTAGCGGCGATGCCCAACAGCGCGCAGATCGGATCCTGAAAGAGACGGAGCTG	2310
Qy	653	-----	653
Db	2311	AGGAAGTGAAGTGTCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2370
Qy	653	-----	653
Db	2371	CTGTATGGGAGAAATGTGAAAATTCACAGTGGCCATCAAAGTGTGTGAGGAAAAACACATCC	2430
Qy	653	-----	653
Db	2431	CCCAAGCCAAACAAGAAATCTTAGAGAAAGCATACGTGATGGCTGGTGTGGGCTCCCCA	2490
Qy	653	-----	653
Db	2491	TATGTCTCCGCCCTTCTGGGCATCTGCCTGACATCCACGCTGCAGCTGGTGACACAGCTT	2550
Qy	653	-----	653
Db	2551	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAAACCGCGGACGCTGGGCTCCAG	2610
Qy	653	-----	653
Db	2611	GACCTGCTGAACGTGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG	2670
Qy	653	-----	653
Db	2671	CTCGTACACAGGACATTGGCCGCTCGGAAGCTGTGTCTCAAGAGTCCCAACCATGTCAAA	2730
Qy	653	-----	653
Db	2731	ATTACAGACTTCGGGTGGCTCGGCTCGTGCACATTTGACGAGACAGAGTACCATGCAGAT	2790
Qy	653	-----	653
Db	2791	GGGGCAAGTGGCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGCTTCACC	2850
Qy	653	-----	653
Db	2851	CACCAGAGTGTGTGTGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2910
Qy	653	-----	653
Db	2911	AAACCTTACGATGGGATCCCGACCCCGGGAGATCCCTGACCTGCTGGAAAAAGGGGACGG	2970
Qy	653	-----	653
Db	2971	CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTGTTGATG	3030
Qy	653	-----	653
Db	3031	ATTGACTCTGAATGTCCGCCAAGATTCCCGGAGTTGGTGTCTGAAATTTCTCCCGCATGGCC	3090
Qy	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	3091	AGGACCCCCAGCGCTTTTGGGTCTATCCAGAAATGAGGACTTGGGCCAGCGCATCCCTTG	3150
Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3151	GACAGCACCTTCTACCGCTCACCTGCTGGAGAGCATGACATGGGGACCTGGTGGATGCT	3210
Qy	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3211	GAGGAGTATCTGGTACCCCAAGCAGAGGCTTCTGTCTCAGACCTTGCCTGGCGCGCTGGG	3270
Qy	705	GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr	724
Db	3271	GGCATGGTCCACCACAGGACCCAGCTCATCTACAGGAGTGGCGGTGGGACCTTGACA	3330
Qy	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
Db	3331	CTAGGCTGGAGCCCTCTGAAGAGAGGAGGCCCCAGGCTCTCCACTGGCACCTCCGAAGGG	3390
Qy	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer	764
Db	3391	GCTGGCTCCGATGATTTTGTGTGACCTGGGAATGGGGCAGCCAAAGGGCTTGCAAAGC	3450

```

QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db 3451 CTCGCCACACATGACCCAGCCCTCTACAGCGGTACAGTGAGGACCCACACAGTACCCCTG 3510
QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3511 CCCTCTGAGCTGATGGCTAGTTCGCCCTCTGACCTGCAGCCCGCCAGCCTGAATATATGTG 3570
QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3571 AACACGCCAGATGTTCCGCCCGCCAGCCCTTCGCCCGCCAGGAGGCCCTCTGCCCTGCTGCC 3630
QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3631 CGACCTGTGTGGTGCACCTCTGAAAGGGCCAGACCTCTCTCCACAGGAAGAAATGGGGTC 3690
QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3691 GTCAAAGAGCTGTTTGGCCCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACACCCCGAG 3750
QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3751 GGAGGAGTGCCTTCAGCCCGCCAGCCCTCTCTGCTTCAGCCCGACGCTTCGACAAACCTC 3810
QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3811 TATTACTGGGACGAGGACCCAGAGCGGGGGGCTCCACCCAGCACCTTCAAAGGGACA 3870
QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3871 CCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTGCCAGTG 3915

RESULT 9
US-09-877-177A-11
; Sequence 11, Application US/0987177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-11

Alignment Scores:
Pred. No.: 0 Length: 4530
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservative: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 4 Gaps: 1

SEQ6 (1-919) x US-09-877-177A-11 (1-4530)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGGAGCTGGCGCCTTGTGGCGCTGGGGCTCCCTCTCGCCCTCTTGGCCCGCGAGCC 210
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 GCGAGCACCAAGTGTGACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 270
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCACCTGGACATGCTCCGCCACCTCTACAGGGGCTGCCAGGTGGTCAGGGGAAACCTG 330

```

```

QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 331 GAACCTACCTACCTGCCACCAATGCCAGCCTGTCTCTCCCTGCAGGATATCCAGAGGTG 390
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTACGTCGCTCATCGCTCACAAACCAAGTGAGGCAGGTCCCACTCCAGAGGTGGG 450
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGCCAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGCTGCTAGCAATAGA 510
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCGCTGAACAATACCAACCCCTGCTCAGAGGGCTCCCCAGGAGGCTCGCGGAGCTG 570
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 571 CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCGAG 630
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGCTACCAAGGACCATTTTGTGAAGAGGACATCTTCCACAAGAACACACAGTGGCT 690
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 750
QY 201 GlySerArgCysTyrGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTCTCAGAGCTGTCAGCGCACTCTCTGT 810
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 811 GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCACTGCTGCTGCATGAGCAGTGT 870
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 871 GCTGCCGCTGCAGCGGCCCAAGCACTGCTGACTGCTGCTGCCCTCCACTTCAACCCAC 930
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 931 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACACACAGACACGTTTGAG 990
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 991 TCCATGCCCAATCCCGAGGGCGGTATACATTGGCGCCAGCTGTGTGACTGCCTGTCTCC 1050
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1051 TACAACACTACCTTTCTACGAGCGTGGGATCCTGCACCCCTGCTGCCCTCCACCAACAA 1110
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1111 GAGGTGACAGCAGAGAGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCTGTGCCCGA 1170
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1171 GTGTGCTATGGTCTGGGCATGAGCACCTGGCAGAGGTGAGGGCAGTACCAGTGCCTCAAT 1230
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1231 ATCCAGGAGTGTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATCTTGTCCGCGAGAG 1290
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe 400
Db 1291 TTTGTATGGGAGCCCAAGCCTCCAAACACTGCCCGCTCCAGCCAGACAGCTCCCAAGTGT 1350
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1351 GAGACTCTGGAGAGATCACAGTTTACCTATACATCTCAGCATGTGCCGACAGCTGCCT 1410

```

Qy 421 AspLeuSerValPheGlnAsnLeuValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 Db 1411 GACCTCAGCGTCTCCAGAACCTCAAGTAATCCGGGACGAATTCGCACAATGGCGCC 1470  
 Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 Db 1471 TACTCGGTGACCCCTGCAAGGGCTGGCATCAGCTGCTGGGGCTCGCGCTCACTGAGGAA 1530  
 Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480  
 Db 1531 CTGGGCAGTGACCTGGCCCTCATCCACATACACCCACCTCTCGCTCGTCACACCGTG 1590  
 Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 Db 1591 CCCTGGGACCAAGCTCTTTGGGAACCGCACCAAGCTCTGCTCCACACTGGCCAAACCGGCCA 1650  
 Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 Db 1651 GAGGACGAGTGTGGGGCGAGGGCTGGCCCTGCCACCAAGCTGTGCGCCCGAGGCACTGC 1710  
 Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 Db 1711 TGGGGTCAGGGCCCCCAGCTGTGTCACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1770  
 Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTrpValAlaAsnAlaArgHisCys 560  
 Db 1771 GTGGAGGAATGCCAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT 1830  
 Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 Db 1831 TTGGCGTGCCACCCCTGAGTGTACGCCCCAGAAATGGCTCAGTGACCTGTTTTGGACCGGAG 1890  
 Qy 581 AlaAspGlnCysValAlaCysAlaHisTrpLysAspProPheCysValAlaAlaArgCys 600  
 Db 1891 GCTGACCAAGTGTGGCGCTGTGCCCACTATAAGACCCCTCCCTTCTCGGTGGCCCGCTGC 1950  
 Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
 Db 1951 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG 2010  
 Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640  
 Db 2011 GGCCGATGCCAGCCTTGCCCATCACTGCACCCACCTCTGTGTGGAGCTGGATGACAAG 2070  
 Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
 Db 2071 GGCTGCCCGCGAGCAGAGCCAGCCCTCTGACGTCCATCGTCTCTGCGGTGGTGGC 2130  
 Qy 653 2131  
 Db 2131 ATCTCTGCTGCTGCTGTGGGTGGTCTTTGGGATCCTCATCAAGCGAGCGCAGCAG 2190  
 Qy 653 2191  
 Db 2191 AAGATCCGGAAGTACACGATCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTG 2250  
 Qy 653 2251  
 Db 2251 ACACCTAGCGGAGGATGCCCAACCGCGAGATGCGGATCTCTGAAAGAGACGGAGCTG 2310  
 Qy 653 2311  
 Db 2311 AGGAAGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAGGGCATCTGGATC 2370  
 Qy 653 2371  
 Db 2371 CCTGATGGGAGAAATGTGAAAATTCAGTGGCCATCAAAAGTTGTGAGGGAAAAACACATCC 2430  
 Qy 653 2431  
 Db 2431 CCCAAAGCCAAAGAAATCTTTAGACGAAGCATACGTGATGGTGTGGCTGCCCA 2490  
 Qy 653

Db 2491 TATCTCTCCCGCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2550  
 Qy 653 2551  
 Db 2551 ATGCCCTATGCTGCTCTTTAGACCATGTCCGGGAAAAACCGGACGCTGGGCTCCAG 2610  
 Qy 653 2611  
 Db 2611 GACCTGCTGAACCTGCTATTCAGATTTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG 2670  
 Qy 653 2671  
 Db 2671 CTGCTACACAGGAGCTTTGGCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATCTCAAA 2730  
 Qy 653 2731  
 Db 2731 ATTACAGACTTCGGGCTGGCTCGCTGCTGGACATTTGACGAGACAGATACCATGCAGAT 2790  
 Qy 653 2791  
 Db 2791 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTTCTCCGCCCGCGGTTCACC 2850  
 Qy 653 2851  
 Db 2851 CACCAGAGTGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGGCC 2910  
 Qy 653 2911  
 Db 2911 AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGGAGCGG 2970  
 Qy 653 2971  
 Db 2971 CTGCCCCAGCCCCCATCTGCACCATTTGATGCTACATGATCATGGTCAAAATGTTGGATG 3030  
 Qy 653 3031  
 Db 3031 ATTGACTCTGAATGTGCGCAAGATTCCGGGAGTTGGTGTCTGAATTTCTCCCGCATGSCC 3090  
 Qy 654 3091  
 Db 3091 AGGGACCCCCAGCGCTTTGTGGTCAATCCAGATCAGGACTTTGGGCCCGCAGCGCTCCCTTG 3150  
 Qy 665 3151  
 Db 3151 GACAGCACTTCTACCCCTCACTGCTGGAGGACCATGACATGGGGGACCTGCTGGATGCT 3210  
 Qy 685 3211  
 Db 3211 GAGGAGTATCTGGTACCCAGCAGCGGCTTCTTCTGTCACAGACCTGCCCCGGCGCTGGG 3270  
 Qy 705 3271  
 Db 3271 GGCATGTCTCCACACAGCAGCCGAGCTCATCTACAGGAGTGGCGGTGGGAGCTGCACA 3330  
 Qy 725 3331  
 Db 3331 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCCCAGGTCTCCACTGGCACCTCCGAAGGG 3390  
 Qy 745 3391  
 Db 3391 GCTGGCTCCGATGATTTGATGGTACCTGGGAATGGGCGAGCAAGGGGCTCCAAAGC 3450  
 Qy 765 3451  
 Db 3451 CTCCCCACATGACCCAGCCCTCTACAGCGGTACAGTGGAGGACCCACAGTACCCTG 3510  
 Qy 785 3511  
 Db 3511 CCCTCTGAGACTGATGGCTACGTTGCCCCCTGACCTGACCTGCGAGCCCCAGCTGAAATGTG 3570  
 Qy 805 804  
 Db 804 824

Db 3571 AACACGCGCATGTTTCGGCCCCAGCCCCCTTCGCCCGAGAGGGCCCTCTGCCTGTGTGCC 3630  
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
 Db 3631 CGACCTGCTGGTGGCACTCTGGAAGGGCCAGACTCTCTCCCGAGGAGAAATGGGGTC 3690  
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
 Db 3691 GTCAAGACGCTTTTTCCTTGGGGTGGCGGTGAGAACCCCGAGTACTTGACACCCAG 3750  
 QY 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884  
 Db 3751 GGAGGAGCTGCCCTCAGCCCCACCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3810  
 QY 885 TyrTyrTyrAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
 Db 3811 TATTACTGGACGACCCAGCCAGAGCGGGGGCTCCACCCAGACCTTCAAAAGGAGCA 3870  
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
 Db 3871 CCTACGGCAGAGAACCCAGTAGTACCTGGGTCTGGAGTGCAGTG 3915

## RESULT 10

US-08-229-515A-14  
 ; Sequence 14, Application US/08229515A  
 ; Patent No. 5518885  
 ; GENERAL INFORMATION:  
 ; APPLICANT: RAZIUDIN  
 ; APPLICANT: SARKAR, FAZLUL H  
 ; TITLE OF INVENTION: ERB2 PROMOTER BINDING PROTEIN IN  
 ; TITLE OF INVENTION: NEOPlastic DISEASE  
 ; NUMBER OF SEQUENCES: 19  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: NEEDLE & ROSENBERG PC  
 ; STREET: 127 Peachtree Street, Suite 1200  
 ; CITY: Atlanta  
 ; STATE: Georgia  
 ; COUNTRY: usa  
 ; ZIP: 30303  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/229,515A  
 ; FILING DATE: 19 APR 1994  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: PERRYMAN, DAVID G  
 ; REGISTRATION NUMBER: 33,438  
 ; REFERENCE/DOCKET NUMBER: 1414.608  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 404-688-0770  
 ; TELEFAX: 404-688-9880  
 ; INFORMATION FOR SEQ ID NO: 14:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3955 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-229-515A-14

Alignment Scores:  
 Pred. No.: 9,44e-287 Length: 3955  
 Score: 4134.00 Matches: 781  
 Percent Similarity: 65.47% Conservative: 42  
 Best Local Similarity: 62.13% Mismatches: 96  
 Query Match: 81.41% Indels: 338  
 DB: 1 Gaps: 3

SEQ6 (1-919) x US-08-229-515A-14 (1-3955)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20  
 Db 26 ATGGAGCTGGCGGCTGGTGGCGCTGGGTTCTCTCTCGCCCTCTCTCGCCCGGAAATC 85  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 Db 86 GCGGGACCCCAAGTGTATCCGGCACAGACATGAAGTTGGGCTCTCTGCCACTCTCTGAG 145  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlyAsnLeu 60  
 Db 146 ACCCACTCGACATGCTCGCCACCTGTACCAAGGGCTGTCTAGTAGTGTAGGAGCACTTG 205  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
 Db 206 GAGCTTACCTACGTGCTGCAATGCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 265  
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 Db 266 CAGGGTTACATGCTCATCGCTCAACAACAGGTGAAGCGGTCCCACTGTCAAGAGCTGCGC 325  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 Db 326 ATCGTGAGAGGGACCCAGCTCTTTGAGGACAAGTATGCCCTGCTGTCTAGACAACCGA 385  
 QY 121 AspProLeuAsnAsnThrThrProValThr---GlyAlaSerProGlyGlyLeuArgGlu 139  
 Db 386 GATCCTCAGGACAATGTCGCGGCTCCACCCAGGAGAACCCAGAGGGGTGCGGGAG 445  
 QY 140 LeuGlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnPro 159  
 Db 446 CTGCAGCTTCGAAGTCTCAGAGATCTTGAAGGGAGGAGTTTGTATCGGTGGGAACCT 505  
 QY 160 GlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeu 179  
 Db 506 CAGCTCTGTACCAGGACATGTTTGTGGAAGGAGCTCTTCGCAAGATAATACCAACTG 565  
 QY 180 AlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCys 199  
 Db 566 GCTCCTGTGATATAGACACCAATCTTCCGGGCTGTCCACTGTCTGCTGCCCCGCTGC 625  
 QY 200 LysGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrVal 219  
 Db 626 AAAGACAATCACTGTTGGGTGAGTCCGGAAGACTGTCAAGTCTGACTGGCCACCATC 685  
 QY 220 CysAlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGln 239  
 Db 686 TGTACCAGTGGTTGTGCCCGGTGCAAGGCGGCTGCCCACTGCTGCTCCATGAGCAG 745  
 QY 240 CysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsn 259  
 Db 746 TGTGCCGAGGCTGCACGGGCCCCAAGCATTTCTGACTGCTGCTGCTGCTGCTGCTGCT 805  
 QY 260 HisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPhe 279  
 Db 806 CATAGTGGTATCTGTGAGTGTGCTGCCAGCCCTGTCACTTACCAACACAGACACTTT 865  
 QY 280 GluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCys 299  
 Db 866 GAGTCCATGCACAACCTTGAGGTCGTACACCTTTGGTCCAGCTGCTGCTGCTGCTGCT 925  
 QY 300 ProTyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsn 319  
 Db 926 CCTACAACCTACCTGTCTAGGAAAGTGGATCTCTGCACTCTGCTGCTGCTGCTGCTGCT 985  
 QY 320 GlnGluValThrAlaGluAspGlyThrGlnArgCysGlnLysCysSerLysProCysAla 339  
 Db 986 CAAGAGGTCACTGAGTGGAGGACACAGCGTGTGAGAAATGCAAGAGCCCTGGCT 1045  
 QY 340 ArgValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAla 359  
 Db 1046 CGAGTGTCTATGTTCTGGGCATGAGACCTTTCGAGGGGCGAGGCCCATCACCAGTGAC 1105  
 QY 360 AsnIleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGlu 379





QY 763 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 782  
 Db 3326 CAGAGCCTCTCCACATGACCTCAGCCCTCTACAGCGGTACAGGAGGACCCACATTA 3385  
 QY 783 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGluProGlu 802  
 Db 3386 CCTCTGCCCGCCGAGACTGATGCTATGTTGCTGCCCTGGCTGCGCCGAGCCCGAG 3445  
 QY 803 TyrValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuPro 822  
 Db 3446 TATGTGAACCAATCAGAGGTTTCCAGCTCAGCCCTCTTAAACCCAGAGGGTCTCTGCT 3505  
 QY 823 AlaAlaArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsn 842  
 Db 3506 CCTGTCCGCGCTCTGCTACTCTAGAAAGCCAGACTCTCTCTCTCTGGGAGAAAT 3565  
 QY 843 GlyValValAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 862  
 Db 3566 GGGGTGTCAAGAGCGTTTTTGCCTTCGGGGGTGCTGTGGAGAACCTGAATAGTA 3625  
 QY 863 ProGlnGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 882  
 Db 3626 CCGAGAGAGGCGCTCTCTCCGCCCCACCTCTCTCTGCTTCCAGCCCGCTTGTGAC 3685  
 QY 883 AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 902  
 Db 3686 AACCTCTATTACTGGGACCAAGACTCATCGGAGAGCGGCGCTCCACCAAGTAACCTTGA 3745  
 QY 903 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
 Db 3746 GGGAGCCCGCTGACAGAACCTGAGTACCTAGGCTAGGCTGGATGATACCTGTA 3796

RESULT 11

US-08-645-865-14  
 ; Sequence 14, Application US/08645865  
 ; Patent No. 5654406  
 ; GENERAL INFORMATION:  
 ; APPLICANT: RAZIUDIN  
 ; APPLICANT: SARKAR, FAZLUL H  
 ; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN  
 ; TITLE OF INVENTION: NEOPLASTIC DISEASE  
 ; NUMBER OF SEQUENCES: 19  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: NEEDLE & ROSENBERG PC  
 ; STREET: 127 Peachtree Street, Suite 1200  
 ; CITY: Atlanta  
 ; STATE: Georgia  
 ; COUNTRY: usa  
 ; ZIP: 30303  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/645,865  
 ; FILING DATE: 14 MAY 1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: PERRYMAN, DAVID G  
 ; REGISTRATION NUMBER: 33,438  
 ; REFERENCE/DOCKET NUMBER: 1414.608  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 404-688-0770  
 ; TELEFAX: 404-688-9880  
 ; INFORMATION FOR SEQ ID NO: 14:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3955 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-645-865-14

Alignment Scores:  
 Pred. No.: 9,44e-287 Length: 3955  
 Score: 4134.00 Matches: 781  
 Percent Similarity: 65.4% Conservative: 42  
 Best Local Similarity: 62.1% Mismatches: 96  
 Query Match: 81.4% Indels: 338  
 DB: 1 Gaps: 3  
 SQ06 (1-919) x US-08-645-865-14 (1-3955)  
 QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuLeuLeuProGlyAla 20  
 Db 26 ATGGAGCTGGCGGCTGGTGGCGTGGGGTTCCTCCTCGCCTCTCTCCGCCCGGAAATC 85  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProGlu 40  
 Db 86 GCGGCGACCAAGTGTACCGGCACAGACATGAAGTTGCGGCTCCCTGCCAGTCTCTGAG 145  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
 Db 146 ACCCACTGGACATGCTCGGCCACCTGTACACAGGCTGTACAGTAGTCAGGGCAACTTG 205  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80  
 Db 206 GAGCTTACCTAGCTGCTGCCAATGCCAGCCTCTCATTCCTGCAGGACATCCAGGAAGTT 265  
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeu 100  
 Db 266 CAGGGTTACATCATCGCTCACAAACAGGTGAAGCGCGTCCCACTGCAAGGCTGCGC 325  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 Db 326 ATCGTGAGAGGACCCAGCTCTTTGAGGACAGTAGTATGCCCTGGCTGTCTAGACAAAC 385  
 QY 121 AspProLeuAsnAsnThrThrProValThr---GlyAlaSerProGlyGlyLeuArgGlu 139  
 Db 386 GATCCTCAGGACAATGTGCGCCCTCCACCCAGGAGGAGGAGGAGGAGGAGGAGGAG 445  
 QY 140 LeuGlnLeuArgSerLeuThrGlnLeuLeuLysGlyValLeuIleGlnArgAsnPro 159  
 Db 446 CTGCAGCTTCGAAGTCTCACAGATCTCTGAAGGAGGAGGAGGAGGAGGAGGAGGAG 505  
 QY 160 GlnLeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeu 179  
 Db 506 CAGCTCTGCTACACGAGACATGCTTTTGTGGAAGAGCTTCTCCGCAAGAAATAACCAACTG 565  
 QY 180 AlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCys 199  
 Db 566 GCTCCTGTGATATAGACACCAATGTTCCCGGCGCTGTCCACCTTGTGCCCGCGCTGC 625  
 QY 200 LysGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrVal 219  
 Db 626 AAAGACATCATCTGTTGGGGTGAGAGTCCGGAAGACTGTGACATCTTGTACTGGCACCAT 685  
 QY 220 CysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGluGln 239  
 Db 686 TGTACCACTGTTGTCGCGGTGCAAGGCGCGGCTGCCACTGTGCTGCTGATGAGCAG 745  
 QY 240 CysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsn 259  
 Db 746 TGTGCCGAGGCTGCACGGGCCCAAGCATCTGACTGCTGCTGCTGCTGCTGCTGCTGCT 805  
 QY 260 HisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPhe 279  
 Db 806 CATAGTGTATCTGTGAGCTGCACCTGCCAGCCCTCGTACCTACACACACACACACCTTT 865  
 QY 280 GluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCys 299  
 Db 866 GAGTCCATGCACACCCCTGAGGGTGCCTACACTTTGGTGGCCAGCTGCTGCTGCTGCTG 925  
 QY 300 ProTyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsn 319  
 Db 925

Db	926	CCCTACAACTACCTGTCTACCGAAGTGGGATCCCTGCACCTCGGTGTGTCTCCCGCAATAAC	985
Qy	320	GlnGluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAla	339
Db	986	CAAGAGGTCACAGCTGAGGACGGAACACAGCGTTGTGAGAAATGCAGCAACCCCTGTGCT	1045
Qy	340	ArgValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAla	359
Db	1046	CGAGTGTCTATGCTGTGGGATGAGCACCTTCAGGGGGGAGGGCCATCACCAGTGAC	1105
Qy	360	AsnIleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGlu	379
Db	1106	ANTGTCAGGAGTTGGCTGCAAGAAGATCTTGGAGCCCTGGCATTTTTCGGCGAG	1165
Qy	380	SerPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnVal	399
Db	1166	AGCTTTGATGGGACCCCTCTCCGGCATTCCTCCGCTGAGGCCCTGAGCAGCTCCNAGTG	1225
Qy	400	PheGluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeu	419
Db	1226	TTCCGAACCCCTGGAGGAGATCACAGTTACCTGTACATCTCAGCATGCCAGACAGTCTC	1285
Qy	420	ProAspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGly	439
Db	1286	CGTGACCTCAGTGTCTCCAGAACCTTCGAATCATTCGGGACGCGATCTCCACGATGGC	1345
Qy	440	AlaTyrSerLeuThrLeuGlnClyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArg	459
Db	1346	CGGTACTATTGACATGCAAGGCTGGGGATCCACTCGCTGGGGCTGCGCTCACTCGGG	1405
Qy	460	GluLeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThr	479
Db	1406	GAGCTGGCAGTGGTGGCTCTGATTACCGCAACGCCCATCTCTGCTTTGTACACACT	1465
Qy	480	ValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArg	499
Db	1466	GTACCTTGGGACCAAGCTCTCCGGAACCCACATCAGGCCCTGCTCCACAGTGGGAACCGG	1525
Qy	500	ProGluAspGlu---CysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGly	518
Db	1526	CCGGAAGAGGACTTGTGGCTTCGAGCGGCTGGTCTGTACTCACTGTGTGCCACCGG	1585
Qy	519	HisCysTrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGln	538
Db	1586	CACCTGCTGGGGCCAGGGCCACCACAGTGTCTCAACTGCAGTCACTTCTTCGGGGCCAG	1645
Qy	539	GluCysValGluGluCysArgValLeuGlnClyLeuProArgGluTyrValAsnAlaArg	558
Db	1646	GAGTGTGTGGAGAGTGCCGAGTATGGAAGGGCTCCCCGGGAGTATGTGAGTGACAAG	1705
Qy	559	HisCysLeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGly	578
Db	1706	CGCTGTCTGCCGTGTACCCCGAGTGTGAGCTCAAAACAGCTCAGAGACCTGCTTTTGA	1765
Qy	579	ProGluAlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAla	598
Db	1766	TCGGAGGCTGATCAGTGTGACAGCTGCGCCCACTACAAGGACTCGCTCTCTGTGTGCT	1825
Qy	599	ArgCysProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAsp	618
Db	1826	CGCTGCCCCAGTGTGTGAACCGGACCTCTCTACATGCCCCATCTGGAAGTACCCGGAT	1885
Qy	619	GluGluGlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAsp	638
Db	1886	GAGAGGGCATATCCAGCCGCTGCCCATCACTACATGCAACCACTCCTGTGTGGATCTGGAT	1945
Qy	639	AspGlyCysProAlaGluGlnArgAlaSerProLeuThr-----	652
Db	1946	GAACGAGGCTGCCAGCAGACAGAGCCAGCCCGGTGACATTCATTCATTGCAACTGTA	2005
Qy	652	-----	652
Db	2006	GAGGGCTGCTGTGTTCTCTGATCTTAGTGGTGGTCTGTGGAATCCTTAATCAACGAAGG	2065
Qy	652	-----	652
Db	2066	AGACAGAAGATCCCGAAGTATACGATCGGTAGGCTGCTGCAGGAACACTGACTTAGTGAG	2125
Qy	652	-----	652
Db	2126	CCGCTGACGCCAGCGGAGCAATGCCAACCAAGGCTCAGATCGGATCCTTAAAGAGACG	2185
Qy	652	-----	652
Db	2186	GAGCTAAGGAAGTGAAGTCTTGGATCAGGAGCTTTTGGCACTGTCTACAAGGGCATC	2245
Qy	652	-----	652
Db	2246	TGGATCCAGATGGGAGAAATGTGAAAATCCCCGTGGCTATCAAGGTGTTGAGAGAAAAC	2305
Qy	652	-----	652
Db	2306	ACATCTCCTAAAGCCAAACAAAGAAATTCATAGATGAAGCTATGTATGGTGGTGGGT	2365
Qy	652	-----	652
Db	2366	TCTCCGTATGTCTCCCGCTCTCTGGGCATCTGCTGACATCCACAGTACAGCTGGTGACA	2425
Qy	652	-----	652
Db	2426	CAGCTTATGCCCTACGGCTGCCCTTCTGGACCATGTCCGAGAACACCGAGGTCGCCATAGGC	2485
Qy	652	-----	652
Db	2486	TCCCAGGACCTGCTCAACTGGTGTTCAGATTGCCAAGGGGATGAGCTACCTGGAGGAC	2545
Qy	652	-----	652
Db	2546	GTGCGGCTGTACACAGGAGCCTGGCTGCCCGGAATGTGCTAGTCAAGAGTCCCAACCCAC	2605
Qy	652	-----	652
Db	2606	GTCAAGATTACAGATTTCCGGGCTGGCTCGGCTGGACATTTGATGACAGAGTACCAT	2665
Qy	652	-----	652
Db	2666	GCAGATGGGGCAAGGTGCCCATCAAAATGGATGGCATTTGGAATCTATTCTCAGACGCGG	2725
Qy	652	-----	652
Db	2726	TTCAACCATCAGAGTGTGTGTGGAGCTATGGAGTGTGTGGGAGCTGATGACTTTT	2785
Qy	652	-----	652
Db	2786	GGGGCCAAACCTTACGATGGAAATCCAGAGCCCGGAGATCCCTGATTTGCTGGAGAGGGA	2845
Qy	652	-----	652
Db	2846	GAAGCGCTACCTCAGCCCTCCAATCTGCACCATTTGATGTCTACATGATTTGCTCAAAATGT	2905
Qy	652	-----	652
Db	2906	TGGATGATTGACTCTGAATGTGCGCCGAGATTCGGGAGTGTGGTGTGAGAAATTTTCACGT	2965
Qy	653	-----SerGlnAsnGluAspLeuGlyProAlaSer	662
Db	2966	ATGGCGAGGAGACCCCGAGCGTTTGTGGTTCATCAGAAACGAGGACTTTGGGCCCATCCAGC	3025
Qy	663	ProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuVal	682
Db	3026	CCCATGGACAGTACCTTCTACCGTTCACTGCTGGAAGATGATGACATGGGTGACCTGGTA	3085
Qy	683	AspAlaGluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGly	702
Db	3086	GAGCTGAAGAGTATCTGGTGGCCCGCAGCAGGAGTTCTTCCCGCGGACCTTACCCAGGC	3145

```

QY 703 AlaGlyGlyMetValHisArgHisArgSerSerThrArgSerGlyGlyGlyAsp 722
Db 3146 ACTGGGAGCAGACCCATAGAAGCCACCCAGCTCCACCAGGAGTGGAGTGGTGGAG 3205
QY 723 LeuThrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSer 742
Db 3206 CTGACACTGGGCTGGAGCCCTCGGAAGAAGGCGCCCGCCAGATCTCCACTGGCTCCCTGG 3265
QY 743 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeu 762
Db 3266 GAAGGGCTGGCTCCGATGTTGATGTGACCTGGCAATGGGGTACCAAGAGGGCTG 3325
QY 763 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 782
Db 3326 CAGAGCTCTCTCCACATGACCTACGCTCTACAGCGGTACAGCGAGGACCCACATTA 3385
QY 783 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGlu 802
Db 3386 CCTCTGCCCGCCGAGACTGATGGCTATGTGTGCTCCCTGGCCTGCAGCCCGCCGAG 3445
QY 803 TyrValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuPro 822
Db 3446 TATGTGAACCATCAGAGGTTGAGCTCAGCTCTCTTTAAACCCAGAGGGTCTCTGGCT 3505
QY 823 AlaAlaArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsn 842
Db 3506 CCTGTCCGGCTGCTGGTGTACTCTAGAAAGACCCAGACTCTCTCTCTGGGAAGAT 3565
QY 843 GlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrThr 862
Db 3566 GGGGTGTCAAGACGCTTTTGGCTTCGGGGGTGCTGTGGAGAACCTGAATACTTAGTA 3625
QY 863 ProGlnGlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAsp 882
Db 3626 CCGAGAGAGGACTGCTCCCTCTCGGCCACCCCTCTCTCTGCTTCCAGCCAGCCCTTGGAC 3685
QY 883 AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 902
Db 3686 AACCTCTATTACTGGGACCAAGAACTCATCGGAGCAGGGGCTCCACCAAGTAACCTTTGAA 3745
QY 903 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3746 GGGACCCCACTCAGAGAACCCGTGAGTACCTAGGCTGGATGATACCTGTA 3796

RESULT 12
US-09-146-283-3
; Sequence 3, Application US/09146283
; Patent No. 5976546
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702

```

```

; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
; US-09-146-283-3

Alignment Scores:
Pred. No.: 4,4e-251 Length: 2385
Score: 3632.00 Matches: 659
Percent Similarity: 98.65% Conservative: 0
Best Local Similarity: 98.65% Mismatches: 5
Query Match: 71.52% Indels: 4
DB: 2 Gaps: 2

SEQ6 (1-919) x US-09-146-283-3 (1-2385)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 11 ATGGAGCTGGCGGCTTGTGGCGCTGGGGCTCTCTCGCTCTTGGCCCCGGAGCC 70
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 71 GCGAGACCAAGTGTGACCGGCACAGACATAGAGCTCGGCTCCCTGCCAGTCCCGAG 130
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 131 ACCACCTGGACATGCTCGCCACCTCTACCAGGGCTGCAGGTGGTGCAGGAAACCTG 190
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnGluVal 80
Db 191 GAACCTACCTACCTGCCACCAATGCCAGCTGTCTTCTCCTCAGGATATCCAGGAGTG 250
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 251 CAGGGCTAGCTGCTCATCGCTCACACCAAGTGGAGCAGCTCCACTGCAGAGGCTGCGG 310
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 311 ATTTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGCCCTGTCTAGACAAATGA 370
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 371 GACCCGCTGAACAATACCACTGCTGCAGAGGGCTTCCCGAGGAGCTTGGGGAGCTG 430
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 431 CAGCTTCGAGGCTCACAGAGATCTTGAAGAGGGGTCTTGTATCCAGCGGAAACCCCGAG 490
QY 161 LeuCysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnGlnLeuAla 180
Db 491 CTCTGCTACCCAGGACACGATTTTGTGAAGGACATCTTCCACAGAAACCAACAGCTGGCT 550
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 551 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAG 610
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 611 GGCTCCCGCTGCTGGGAGAGAGTCTGAGGATTTGTACAGGCTGACGCGCAGCTGTCTGT 670
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGluGlnCys 240

```

```
Db 671 GCGGGTGGCTGTGCCCCCTGCAAGGGGCCACTGCCACTGCTGCCATGAGCAGTGT 730
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 731 GCTGCCGGCTGACAGGGGCCCAAGCACTCTGACTGCGCTGCCCTCCACTTCAAGCAC 790
QY 261 SerGlyLeuCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 791 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACCTACACACAGACACGTTTGA 850
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 851 TCCATGCCCAATCCCGAGGGCCGGTATACATTGCGGCCAGCTGTGTGACTGCTGTGCC 910
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 911 TACAACCTACCTTTCTACGAGCTGGGATCTGCAACCTCTGCTGCCCTGCCACACCAA 970
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 971 GAGGTGACAGCAGAGGATGAACACACAGCGGTGTGAGAGTGCAGAGCCCTGTGCCGA 1030
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1031 GTGTGTATGTGTGGGCATGGAGCACCTTGCAGAGGTGAGGCGAGTTACCAAGTGC 1090
QY 361 IleGlnGluPheAlaGlyCysLysLysLysLysPheGlySerLeuAlaPheLeuProGluSer 380
Db 1091 ATCCAGAGTTCTGCTGGCTGCAAGAAGATCTTTGGGAGCGTGGCATTTCTGCCGGAG 1150
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1151 TTTGATGGGGACCCAGCCTCCACACTGCCCTCCAGCCAGCAGCAGCTCCAAGTGT 1210
QY 401 GluThrLeuGluGluLeuThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1211 GAGACTCTGGAAGAGATCAAGGTTTACCTATACATCTCAGCATGGCGGAGCGCTGCCT 1270
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1271 GACCTCAGCGCTCTCCAGAACCTGCAAGTAAATCCGGGGAGGAATCTGCACAATGGCG 1330
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1331 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCAGTGAGGAA 1390
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1391 CTGGGCGAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTGCACAGGGT 1450
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1451 CCTGGGACAGCTCTTTTGGAAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCA 1510
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1511 GAGGACGAGTGTGTGGGCGAGGGCCTGGCTGCCACCACTGTCGCCCGGAGGCACTGC 1570
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1571 TGGGGTCCAGGGCCCAACCCAGTGTCAACTGCAGCCAGTTCTCTCGGGGCCAGGAGTGC 1630
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1631 GTGGAGAAATGCCGAGTACTGCAAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1690
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1691 TTCCCGTGCACCCCTGAGTGTACAGCCCGCAGAAATGGCTCAGTACCTGTTTGGACCGG 1750
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
```

```
Db 1751 GCTGACACAGTGTGGCCTGTGCCACTATAGGACCCCTCCCTTCTGCTGGCGCCGCTGC 1810
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1811 CCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAAGTTTCCAGATGAGGAG 1870
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1871 GCGCATGCCAGCCTTGCCCCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 1930
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGlyPro 660
Db 1931 GGCTGCCCGCCGAGCAGAGAGCCCTCTGACGTCCCTCGAG-----GCACCC 1981
QY 661 Ala---SerProLeuAspSerThr 667
Db 1982 GCGCGTCCGCCAGCCCGCCAGCACA 2005

RESULT 13
US-08-579-823A-3
: Sequence 3, Application US/08579823A
: Patent No. 6080409
: GENERAL INFORMATION:
: APPLICANT: Laus, Reiner
: APPLICANT: Ruegg, Curtis L.
: APPLICANT: Wu, Hongyu
: TITLE OF INVENTION: Immunostimulatory Composition and Method
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dehlinger & Associates
: STREET: 350 Cambridge Ave. Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/579,823A
: FILING DATE: 03-DEC-1998
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Judge, Linda R.
: REGISTRATION NUMBER: 42,702
: REFERENCE/DOCKET NUMBER: 7636-0010
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-324-0880
: TELEFAX: 650-324-0960
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2385 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: homo sapiens
: INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
: US-08-579-823A-3

Alignment Scores:
Pred. No.: 4.4e-251 Length: 2385
Score: 3632.00 Matches: 659
Percent Similarity: 98.65% Conservative: 0
Best Local Similarity: 98.65% Mismatches: 5
Query Match: 71.52% Indels: 4
DB: 3 Gaps: 2
```



Alignment Scores:			
Pred. No.:	4.4e-251	Length:	2385
Score:	3632.00	Matches:	659
Percent Similarity:	98.65%	Conservative:	0
Best Local Similarity:	98.65%	Mismatches:	5
Query Match:	71.52%	Indels:	71
DB:	3	Gaps:	2
SEQ6 (1-919) x US-09-344-195-3 (1-2385)			
Qy	1	MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla	20
Db	11	ATGGAGCTGGCGGCTTGTGGCGTGGGGGTCTCTCTCGCCCTCTGTGCCCCCCCGAGCC	70
Qy	21	AlaSerThrGlnValCysThrGlyThraspMetLysLeuArgLeuProAlaSerProGlu	40
Db	71	GGAGCACCCCAAGTGTGCACCGGCACACATGAAGCTGGGGTCTCCTGCCAGTCCCGAG	130
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrglnGlyCysGlnValValGlnGlyAsnLeu	60
Db	131	ACCCAGCTGGACATGCTCCGCCACCTCTACCCAGGGCTGCCAGGTGGTGCAGGGAACCTG	190
Qy	61	GluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	191	GAACATCACCTACCTGCCACCAATGCCAGCGTGTCTTCTTGCAGGATATCCAGGAGGTG	250
Qy	81	GlnGlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	251	CAGGGCTACGTGCTCATCGCTCACAACTGAGGAGGTGCCACTGCAGAGGCTGGCG	310

```

1391 CTGGCAGTGGAGTGGCCCTCATCCACCAATAACCCACCTCTGCTTGTGCACACGGTG 1450
481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
1451 CCTGGGACCACTCTTTTGGAAACCGCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1510
501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
1511 GAGGACAGTGTGGGCGAGGGCTGGCTGCCACCACTGTGCCCGCCGAGGGCACTGC 1570
521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
1571 TGGGTGCCAGGGCCCAACCCAGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1630
541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAlaAsnAlaArgHisCys 560
1631 GTGGAGGAATGGGAGTACTGAGAGGGCTCCCGCAGGGAGTATGTGAATGCCAGGCACTGT 1690
561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
1691 TTGGCGTGCACCCCTGAGTGTGAGCCCGCAGATGGCTCAGTGACCTGTTTGGACCGGAG 1750
581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
1751 GCTGACCACTGTGTGGCCCTGTGCCACTATAAGGACCTCCCTTCTGCGTGGCCGCTGC 1810
601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
1811 CCAGCGGTGTGAACCTGACCTCTCTATATGCCCATCTGGAAGTTTCCAGATGAGGAG 1870
621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
1871 GCGCATGCGCAGCTTGGCCCATCACTGCACCCACTCTCTGTGTGACCTGGATGACAG 1930
641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGlyPro 660
1931 GGTGTCGCCCGCGAGCAGAGAGCCAGCCCTCTGACCTGCTCCCTGAG-----GCACCC 1981
661 Ala---SerProLeuAspSerThr 667
1982 GCGCGTCCGCCAGCCCGCCAGCACA 2005

```

```

RESULT 15
US-08-422-108-2
; Sequence 2, Application US/08422108
; Patent No. 6015567
; GENERAL INFORMATION:
; APPLICANT: Hudziak, Robert M.
; APPLICANT: Shepard, H. Michael
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,108
; FILING DATE: 14-Apr-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/355460
; FILING DATE: 13-DEC-1994
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: 08/048346
; FILING DATE: 15-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/354319
; FILING DATE: 19-MAY-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 554C2D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/223-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1872 nucleotides
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-422-108-2

Alignment Scores:
Pred. No.: 1,08e-239 Length: 1872
Score: 3471.00 Matches: 623
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 68.35% Indels: 0
DB: 3 Gaps: 0

SEQ6 (1-919) x US-08-422-108-2 (1-1872)
QY 22 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 41
DB 1 AGACCCCAAGUGUGACCGGCACAGACAUGAGGUGCGGCUCCUCCAGUCCCGAGACC 60
QY 42 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGlu 61
DB 61 CACCUAGACUAGCUGCCGCCACCUACACAGGCGUGGUGGUGGAGGAGGAGGAGGAA 120
QY 62 LeuThrTyrProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnValGln 81
DB 121 CUCACCUACUCCGCCACCAUCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCC 180
QY 82 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 101
DB 181 GGUACGUGUGCAUCGUCACACCAAGUGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
QY 102 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 121
DB 241 GUGCGAGGACCCAGCUCUUGAGGACAAUAGCCUCCUCCUCCUCCUCCUCCUCCUCCUCC 300
QY 122 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlnLeuGln 141
DB 301 CCGUGAACAUAUACCCACCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCC 360
QY 142 LeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 161
DB 361 CUUCGAAGCCUCACAGAGAUUUUAGAGGAGGUGUUGAUCCAGCGGAAACCCCGAGCUC 420
QY 162 CysTyrGlnAspThrIleLeuTrpLysAspIlePheHisIlyAsnAsnGlnLeuAlaLeu 181
DB 421 UGUACACGAGACAGAUUUUUGAGGAGCAUCCUCCACAAAGAACACAGCGUGGUCUC 480
QY 182 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 201
DB 481 ACACUGAUAAGACACCAACCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCC 540
QY 202 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 221
DB 541 UCCCGCUGCGGGGAGAGAUUUGAGGAGUUGUAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 222 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCysAla 241

```



Db 601 GGUGGUCUGCCCGUGCAAGGGCCACUGCCACUGACUGCCCAUGACGAGUGUGCU 660  
 Qy 242 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 261  
 Db 661 GCCGGUCGACGGGCCCAAGCACUCUGACUGCCUGGCCUGCCUCCACUUAACACACAGU 720  
 Qy 262 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 281  
 Db 721 GGCAUCUGUGAGUGCAGUGCCAGCCUGGUCACCUACAAACACAGACACACUUUGAGUCC 780  
 Qy 282 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 301  
 Db 781 AUGCCCAUCCCGAGGGCCGGUAUACAUCUGGCCCCAGCUGUGUGACUGCCUGUCCUAC 840  
 Qy 302 AsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGlnGlu 321  
 Db 841 AACUACCUUUCUACGGAGUGGGAUCCUGCACCUGCUGCCUCCUCCUGCACCAACAGAG 900  
 Qy 322 ValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArgVal 341  
 Db 901 GUGACAGCAGAGGAUGGAACACAGCGGUGUGAGAGUGCAGCAAGCCUGUGCCGAGUG 960  
 Qy 342 CysTyrGlyLeuGlyMetGluHisLeuArgGluValAlaArgAlaValThrSerAlaAsnIle 361  
 Db 961 UGUUAUGGUCUGGGCAUGGAGCACUUGCGAGGAGGAGGAGUUAUACGAGUCCCAUAUC 1020  
 Qy 362 GlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSerPhe 381  
 Db 1021 CAGGAGUUGGUCGUCGCAAGACAUUUUGGAGCCUGGCAUUUCGCCGAGAGCUUU 1080  
 Qy 382 AspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPheGlu 401  
 Db 1081 GAUGGGACCCAGCCUCCACACUGCCCGCUGCCAGCAGCAGCAGCAGCAGCAGCAGCAG 1140  
 Qy 402 ThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuProAsp 421  
 Db 1141 ACUCUGGAAGAGAUACAGGUAUACCUAUACUACUACUACUACUACUACUACUACUACU 1200  
 Qy 422 LeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAlaTyr 441  
 Db 1201 CUCAGCGUCUCCAGAGACCUCCAGUAUACCGGGAGGAGAAUUCUGCACAAUGGGCCUAC 1260  
 Qy 442 SerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGluLeu 461  
 Db 1261 UCGCUGACCCUGCAAGGCGUGGCAUCAGCUGGCGUGGCGUCUCACUGAGGGAACUG 1320  
 Qy 462 GlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrValPro 481  
 Db 1321 GGCAGUGGACUGGGCCCUCAUCCACCAUACACCCACCUCUCUGUCGACACGUGGCC 1380  
 Qy 482 TrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgProGlu 501  
 Db 1381 UGGACACAGCUUUUCCGAAACCCGACCAAGCUCUGCUCACUGCCAAACCGGCCAGAG 1440  
 Qy 502 AspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCysTrp 521  
 Db 1441 GACGAGUGUGGGCGAGGGCCUGGCCUGCCACCGCUGGCGCCCGAGGSCACUGCUGG 1500  
 Qy 522 GlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCysVal 541  
 Db 1501 GGUCCAGGGGCCACCCAGUGUGUACAUCGACGAGUCCUUCGGGGCCAGGAGUGCGUG 1560  
 Qy 542 GluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCysLeu 561  
 Db 1561 GAGGAUUGCCGAGUACUGCAGGGGCCUCCCGAGGAGUAUGUAAUCCAGGCACUCUGUUUG 1620  
 Qy 562 ProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGluAla 581  
 Db 1621 CCGUGCCACCUGAGUGCAGGCCCCAGAAUGGUCUGAGCCUUGUUUGGACCGGAGGCU 1680  
 Qy 582 AspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCysPro 601  
 Db 1681 GACCAGUGUGGGCCUUGGCCCAUAUAAGGACCCUCCUUCUGCGUGGGCCGCGUCC 1740

Qy 602 SerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGluGly 621  
 Db 1741 AGCGGUGUGAAACCGACCUCCUACAUGCCCAUGGAAAGUUUCCAGAUAGAGAGGCGC 1800  
 Qy 622 AlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLysGly 641  
 Db 1801 GCAUGCCAGCCUUGGCCCAUCAACAGUGCACCACUCCUGUGUGGACCUUGAUGACAAGGCGC 1860  
 Qy 642 CysProAlaGlu 645  
 Db 1861 UCCCCCGCCGAG 1872

Search completed: October 16, 2003, 17:20:37  
 Job time : 351.963 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 872.869 Seconds  
(without alignments)  
2842.104 Million cell updates/sec

Title: SEQ6

Perfect score: 5078

Sequence: 1 MELAAACRWGLLLALLPPGA.....TFKGTPTAENPEYGLDVPV 919

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlp  
-Q=/cyn2\_1/USPPO\_spool\_p/HOLLERAN480/runat\_15102003\_131912\_20526/app\_query.fasta\_1.4685  
-DB=N\_Geneseq\_19Jun03 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 -CGNVL\_1\_1758 -runat\_15102003\_131912\_20526 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_19Jun03.\*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4900	96.5	3768	17	AAT40739
2	4900	96.5	3768	20	HER-2/neu oncogene
3	4900	96.5	3768	22	Human HER-2/neu on
4	4900	96.5	3768	24	Human HER-2/neu pr
5	4900	96.5	3768	24	Human HER-2/neu pr
6	4900	96.5	3768	24	Human HER-2/neu CD
7	4900	96.5	3768	24	Human HER-2/neu DN
8	4900	96.5	4472	21	ABK10730
9	4900	96.5	4473	19	ABK14812
10	4900	96.5	4473	20	ABQ76220
11	4900	96.5	4473	24	AAZ31071
12	4892	96.3	4473	24	ABZ34969
13	4892	96.3	3768	21	AAO9455
14	4892	96.3	3768	24	ABZ35744
15	4892	96.3	3768	24	ABX09987
16	4892	96.3	3768	24	AAZ43935
17	4892	96.3	3768	24	AAZ43986
18	4892	96.3	3768	24	ABV78168
19	4892	96.3	3768	24	ABV78168
20	4892	96.3	4530	16	AAZ01585
21	4892	96.3	4530	18	AAZ01585
22	4892	96.3	4530	21	AAZ60815
23	4892	96.3	4530	22	AAZ60815
24	4892	96.3	4530	24	ABZ35012
25	4892	96.3	4530	24	ABV94128
26	4892	96.3	4530	24	ABN85585
27	4892	96.3	4530	24	ABK83918
28	4892	96.3	4530	25	ACC00139
29	4892	96.3	4530	25	ABQ83856
30	4892	96.3	9274	24	ABK14057
31	4892	96.3	9274	24	ABK14057
32	4885	96.2	9274	22	AAF24297
33	4885	95.6	4299	14	AAQ46083
34	4722	93.0	3678	24	ABK86207
35	4583	90.3	3600	21	AAA89736
36	4309.5	84.9	2763	24	ABA92252
37	4309.5	84.9	2781	24	ABA92253
38	4134	81.4	3955	16	AAZ01590
39	4134	81.4	3955	21	AAA89753
40	4125.5	81.2	3771	21	AAA89737
41	4125.5	81.2	3771	22	AAH42210
42	4125.5	81.2	3771	24	ABA92251
43	3632	71.5	2385	18	AAZ72725
44	3525	69.4	2871	21	AAZ50586
45	3422	67.4	1872	11	AAQ06828

ALIGNMENTS

RESULT 1

AAT40739  
ID AAT40739 standard; cDNA; 3768 BP.

XX

AC AAT40739;

XX

DT 01-JAN-1997 (first entry)

XX

DE HER-2/neu oncogene.

XX

KW HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;

KW breast cancer; ovary cancer; colon cancer; lung cancer;

KW prostate cancer; genetic immunisation; tumour; vaccine; vector;

XX ss.

XX Homo sapiens.

OS

XX



QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCAGCTCTTTCGGAACCGGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1500
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACGAGTGTGTGGCGAGGGCTGGCCCTGCCACCAAGCTGTGGCCCGGAGGCACTGC	1560
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCAGGGCCACCAAGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC	1620
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGAATGCCAGTACTGCAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACATGT	1680
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGGCGTGCCACCTGAGTGTACGCCCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG	1740
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrIlysAspProProPheCysValAlaArgCys	600
Db	1741	GCTGACCAAGTGTGGCTGTGCCCACTATAAGGACCTCCCTTCTGCGTGGCCGCTGC	1800
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCCAGCGGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640
Db	1861	GGCGCATGCCAGCCTTGCCCCATCAACTGCACCCACCTCCTGTGTGGACCTGGATGACAAG	1920
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCGAGCAGAGCCAGCCCTCTGACGTCCATCATCTCTCGCGTGGTTGGC	1980
QY	653	-----	653
Db	1981	ATTCTGTGTCGTGGTCTTGGGGTGTCTTTGGGATCCTCATCAAGCGACGCGCAG	2040
QY	653	-----	653
Db	2041	AAGATCCGGAAGTACACGATCGGAGACTGTCTGAGGAACAGGAGCTGGTGGAGCGCGTG	2100
QY	653	-----	653
Db	2101	ACACCTAGCGGAGCGATGCCCAACCGCGCAGATGGGATCTCTGAAAGACGAGAGCTG	2160
QY	653	-----	653
Db	2161	AGGAAGGTGAAGTGTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2220
QY	653	-----	653
Db	2221	CCTGATGGGAGAATGTGAAATTTCCAGTGGCCATCAAAGTGTGAGGGAAAAACACATCC	2280
QY	653	-----	653
Db	2281	CCCAAGCCACAAGAAATCTTAGACGAACATACGTGATGGTGTGGGTGCCCA	2340
QY	653	-----	653
Db	2341	TATGTCTCCGCCCTTCTGGGCATCTGCCCTGACATCCAGGTTGCAGCTGGTGACACAGTT	2400
QY	653	-----	653
Db	2401	ATGCCCTATGCTGCTCTTAGACCATGTCCGGGAAACCGCGACGCGCTGGGCTCCAG	2460
QY	653	-----	653
Db	2461	GACCTGTGAACCTGGTGTATCCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG	2520
QY	653	-----	653
Db	2521	CTCGTACACAGGACTTGGCCGCTCGGAACGCTGCTGCTCAAGAGTCCCAACCATGTCAA	2580
QY	653	-----	653
Db	2581	ATTACAGACTTTCGGGCTGGCTCGGCTGTGGACATTGACGACAGACAGTACCATCAGAT	2640
QY	653	-----	653
Db	2641	GGGGCAAGGTGCCCATCAAGTGGATGCGCTGGAGTCCATTTCTCCGCCGCGGTTACCC	2700
QY	653	-----	653
Db	2701	CACCAGAGTGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC	2760
QY	653	-----	653
Db	2761	AAACCTTACGATGGGATCCAGCCGCGGAGATCCCTGACCTGCTGAAAAGGGGAGCGG	2820
QY	653	-----	653
Db	2821	CTGCCCCAGCCCCCATCTGCACCATTCATGTCTACATGATCGTCAAAATGTTGGATG	2880
QY	653	-----	653
Db	2881	ATTGACTCTGAATGTCGGCCCAAGATTCCGGGAGTTGGTGTCTGAATTTCTCCGCATGGCC	2940
QY	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	2941	AGGGACCCCCAGCGCTTGTGTTCATCCAGAAATGAGGACTTGGGCCCGCAGCTCCCTTG	3000
QY	665	AspSerThrPheTyrArgSerLeuLeuAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACTTCTACCGCTCACTGTGGAGACCATGACATGGGGGAGCTGGTGGATGCT	3060
QY	685	GluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGlyAlaGly	704
Db	3061	GAGGATATCTGGTACCCACAGCGCTTCTTCTCCAGACCTGCCCCCGGCGCTGGG	3120
QY	705	GlyMetValHisArgHisArgSerSerSerThrArgSerGlyGlyAspLeuThr	724
Db	3121	GGCATGTGTCCACACAGCACCGACGCTCATCTACAGAGTGGCGGTGGGACCTGACA	3180
QY	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
Db	3181	CTAGGGCTGGAGCCCTCTGAAGAGAGGCCCCCAGGCTCCACCTGGCACCTCCGAAGGG	3240
QY	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer	764
Db	3241	GCTGGCTCCGATGTTTGTGATGGTGACCTGGGAATGGGGGAGCCCAAGGGCTGCAAGC	3300
QY	765	LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu	784
Db	3301	CTCCCCACATGATACCCCGCCCTTACAGCGGTACAGTGGAGGAGCCACAGTACCCCTG	3360
QY	785	ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal	804
Db	3361	CCCTCTGAGACTGATGGCTAGCTTGGCCCCCTGACCTGCAGCCCCCAGCTGGAATATGTG	3420
QY	805	AsnGlnProAspValArgProGlnProProSerProArgGlyProGlyProLeuProAlaAla	824
Db	3421	AACCCAGCATGTTCCGGCCCCCAGCCCCCTTCGCCCGCAGAGAGGCCCCCTTGCCTGCTGC	3480
QY	825	ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal	844
Db	3481	CGACCTGCTGGTGCACTTGGAAAGGCCCAAGACTCTCTCCCCAGGGAAGATGGGGTC	3540
QY	845	ValLysAspValPheAlaPheGlyValAlaValGluAsnProGluTyrLeuThrProGln	864
Db	3541	GTCAAGACGTTTGTGCTTTGGGGTGGCTGGAGAACCCCGAGTACTTGACACCCCGAC	3600
QY	865	GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu	884

```
|||||
Db 3601 GGAGGAGCTGCCCTCAGCCACCCTCCTCTCCTGCTTCCAGCCAGCCTTCGACAACTC 3660
QY 885 TTTTyrTriAspGlnAspProGluArgGlyAlaProSerThrPheLysGlyThr 904
Db 3661 TATTACTGGACCGACCGACCCAGAGCGGGGGCTCCACCCAGACCTTCAAAGGGACA 3720
QY 905 ProThraGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3721 CCTACGGCAGAGAACCAGAGTACCTGGGTCTGGACGTGCCAGTG 3765

RESULT 2
AA01912
ID AAX01912 standard; DNA; 3768 BP.
XX AC AAX01912;
XX DE
XX DT 21-APR-1999 (first entry)
XX HH Human HER-2/neu oncogene DNA.
XX KW HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
XX KW malignancy; treatment; tumour; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CD 1..3768
XX FT /*tag= a
XX FT /product= "HER-2/neu"
XX FT /note= "oncogene"
XX FT misc_feature 2026..3765
XX FT /*tag= b
XX FT /note= "region which elicits immune response"
XX PN US869445-A.
XX PD 09-FEB-1999.
XX PF 01-APR-1996; 96US-0625101.
XX PR 01-APR-1996; 96US-0625101.
XX PR 17-MAR-1993; 93US-0033644.
XX PR 12-AUG-1993; 93US-0106112.
XX PR 31-MAR-1995; 95US-0414417.
XX (UNIW ) UNIV WASHINGTON.
XX PI Cheever MA, Disis ML;
XX DR WPI; 1999-152835/13.
XX DR P-PSDB; AAW92406.
XX PT Use of HER-2/neu polypeptides - for eliciting an immune response to
XX PT an HER-2/neu associated malignancy, particularly for treating or
XX PT preventing tumours
XX PS Claim 1a; Column 23-32; 26pp; English.
XX CC This sequence encodes the human HER-2/neu oncogene protein. A fragment
XX CC of this protein is used in a method for eliciting or enhancing an immune
XX CC response to HER-2/neu protein. The polypeptide can stimulate T cells and
XX CC B cells to produce an immune response to the HER-2/neu protein. The
XX CC method can be used for immunisation against a malignancy in which the
XX CC HER-2/neu oncogene is associated and in the treatment of an existing
XX CC tumour, or to prevent tumour occurrence or reoccurrence.
XX SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 6.12e-227 Length: 3768
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
```

```
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 20 Gaps: 1

SEQ6 (1-919) x AAX01912 (1-3768)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGGAGCTGGCGGCTTGTGGCGCTGGGGCTCCTCTCGCCCTCTTCCCGCGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGACCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCTCTCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCACCTGGACATGCTCGCCACCTCTACCCAGGGCTGCCAGGTGTGAGGGAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 181 GAACCTCACCTACCTGCCCAATGCCAGCTCTCTCTCTCGAGGATATCCAGGAGGTG 240
QY 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTAGCTGCTCATCGCTCAACACCAAGTGAGGCAGGTCCCTCCTGAGAGGTGCGG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGCGCTGTAGACAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCCTGAACAATACCACTCCCTGTACAGGGGCTCCCCAGAGGCTCGCGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyValLeuLeuGlnArgAsnProGln 160
Db 421 CAGCTTCGAAGCCTCAGAGAGATCTTGAAGGAGGGGTCTTGATCAGCGGAAACCCCG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeuAla 180
Db 481 CTCTGTACCGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACACACAGCTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerPrometCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCGATGTGTAG 600
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGTGTGGGAGAGAGTTCTGAGGATTTGTGAGGCTGACAGGCTGACGGCACTGTGT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCGGTGTGTGTGTCCTGCTGCAAGGGGCGCCACTGCCCACTGCTGCTGCCATGAGCAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTGCCGGCTGCACGGGCCCCAAGCACTCTGACTGCTGCGCTGCTCCACTTCAACCCAC 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGGATCTGTGAGCTGCACCTGCCAGGCCCTGGTCACTACACACACAGACAGCTTTGAG 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCATGCCCAATCCCGAGGGCCGGTATACATTCCGGCGCCAGCTGTGTGACTGCTGTCCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 TACAACCTACTTCTTACGGAGCTGGGATCCTGTGACCCCTCTGCTGCCCTGCACAAACCAA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGAGCCCTGTGCCCCGA 1020
```

Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	Qy	653	-----	653
Db	1021	GTGTGCTATGCTTGGCGATGGAGCACTTGGCAGAGGTGAGGCGATTACAGTGCCCAAT	1080	Db	2101	ACACCTAGCGAGCGATGCCCCAACACGCGCAGATGCGGATCCTCAAGAGACACGGAGCTG	2160
Qy	361	IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer	380	Qy	653	-----	653
Db	1081	ATCCAGGAGTTTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCGGAGAGC	1140	Db	2161	AGGAAGTGTAAGTGCTTGGATCTGGGCTTTTGGCACAGTCTACAAAGGCATCTGGATC	2220
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400	Qy	653	-----	653
Db	1141	TTTGATGGGACCCAGCCTCCACACTGCCCGCTCCAGCCAGACGACTCCAAGTGT	1200	Db	2221	CCTGATGGGAGAGAATGTGAAAATTCAGTGCCCATCAAAGTTGTGAGGAAAAACACATCC	2280
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420	Qy	653	-----	653
Db	1201	GAGACTCGGAAGAGATCACAGTTACCTATACATCTCAGCATGGCCGGACAGCCTGCCT	1260	Db	2281	CCCAAGCCCAACAAGAAATCTTAGACGAACATACGTGATGGCTGGTGTGGGCTCCCCA	2340
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440	Qy	653	-----	653
Db	1261	GACCTCAGCGTCTTCCAGAACTGCAAGTAAATCCGGGACGAATCTGCACAAATGGCGCC	1320	Db	2341	TATGTCTCCGCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2400
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460	Qy	653	-----	653
Db	1321	TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGGAA	1380	Db	2401	ATGCCCTATGGCTGCCCTCTTAGACCATGTCCGGGAAAAACCCGCGACGCCCTGGGCTCCAG	2460
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480	Qy	653	-----	653
Db	1381	CTGGCAGTGGACTGGCCCTCATCCACATACACCCACCTCTGCTTCGTGCACACCGTG	1440	Db	2461	GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGGGG	2520
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500	Qy	653	-----	653
Db	1441	CCCTGGGACCAAGCTCTTTGGGAACCCGACCAAGCTCTGTCTCCACACTGCCAACCCGCCA	1500	Db	2521	CTCGTACACAGGAGCTTGGCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAAA	2580
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520	Qy	653	-----	653
Db	1501	GAGGACGAGTGTGGGGAGGGCTGGCCCTGCCACAGCTGTGCGCCCGAGGACACTGC	1560	Db	2581	ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATGTGACAGACAGAGTACCATGCAGAT	2640
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540	Qy	653	-----	653
Db	1561	TGGGTCACAGGCCCCACCCAGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC	1620	Db	2641	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGTTACCC	2700
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560	Qy	653	-----	653
Db	1621	GTGGAGGAATGCCAGTACTGCAGGGCTCCCCAGGAGTATGTGAATGCCAGCACTGT	1680	Db	2701	CACCAGAGTATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2760
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580	Qy	653	-----	653
Db	1681	TTGCGGTGCCACCTGAGTGTACGCCCAAGTGGCTCAGTGACCTGTTTTGGACCCGAG	1740	Db	2761	AAACCTTACGATGGGATCCACGCCCGGAGATCCTGACCTGTCTGGAAGGGGAGCGG	2820
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrIleAspProProPheCysValAlaArgCys	600	Qy	653	-----	653
Db	1741	GCTGACCAGTGTGTGGCTGTGCCCACTATAAGGACCTCCCTTCTCGGTGGCCCGCTGC	1800	Db	2821	CTGCCCCAGCCGCCCATCTGCACCATTTGATGTCTTACATGATCATGTTGGATG	2880
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620	Qy	653	-----	653
Db	1801	CCCAAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTCGAAGTTTCCAGATGAGGAG	1860	Db	2881	ATTGACTCTGAATGTCCGCCAAGATTCCGGAGTTGGTGTCTGAATTTCTCCCGCATGGCC	2940
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640	Qy	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	1861	GGCGCATGCCAGCCTTGCCCACTCAACTGCACCCACCTCTGTGTGGAGCTGGATGACAAG	1920	Db	2941	AGGGACCCCGAGCCTTTGTGGTCAATCCAGAAAGGACTTGGGCCCCAGCCAGTCCCTTG	3000
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	663	Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuAlaAspAla	684
Db	1921	GGTGCCCCCGGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGGTGGTGGC	1980	Db	3001	GACAGCACTTCTACCCCTCAGCTGTGGAGACCATGACATGGGGGACCTGGTGGATGCT	3060
Qy	653	-----	653	Qy	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	1981	ATTCTGCTGTCTGTTGGTCTTTGGGATCTCTTTGGGATCTCATCAAGCAGCGCACAG	2040	Db	3061	GAGGAGTATCTGTGTACCCAGCAGCGCTTCTTCTCTCCAGACCTGCCCCGGCGCTGG	3120
Qy	653	-----	653	Qy	705	GlyMetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyAspLeuThr	724
Db	2041	AAGATCCGGAAGTACACGATCCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTG	2100	Db	3121	GGCATGTCCACCACAGACCGCAGCTCATCTACAGGAGTGGCGGTGGGACCTGACA	3180
				Qy	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744





Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240	Db	1741	GCTGACCAAGTGTGGGCTGTGCCACTATAAGACCCCTCCCTTCTCGTGGCCGCTGC	1800
Db	661	GCCGGTGGCTGTGCCCGCTGCAAGGGCCACTGCCACTGACTGCTGCCATGAGCAGTGT	720	Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu	620
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260	Db	1801	CCCAGCGGTGAAACCTGACCTCTCTACATGCCCATCTCGGAAGTTTCCAGATGAGAG	1860
Db	721	GCTCCCGGCTGACGGGCCCAAGCACTCTGACTGCGCTGCCCTCCACTTCAACAC	780	Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280	Db	1861	GGCGCATGCCAGCCTTGCCCATCACTGCACCCACCTCTGTGTGGACCTGGATGACAAG	1920
Db	781	AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTACAACACAGACACGTTTGA	840	Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300	Db	1921	GGCTGCCCGCCGAG	1980
Db	841	TCCATGCCCAATCCGAGGCGCGGTATACATTCGCGCCAGCTGTGTGACTGCGCTGCC	900	Qy	653	-----	653
Qy	301	TyrAsnThrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320	Db	1981	ATTCTGCTGGTCTGGGCTTGGGGGTGGTCTTTGGGATCTCATCAAGCGAGCGGAGCAG	2040
Db	901	TACAACCTACCTTTCTCGGAGGTGGGATCTCTGCACCTCTGCTGCCGCCCTGCACAACAA	960	Qy	653	-----	653
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340	Db	2041	AAGATCCGGAAGTACACGATCGCGAGACTGCTGCAGGAAACGAGCTGGTGGAGCCGCTG	2100
Db	961	GAGGTGACAGAGAGATGGAACACACAGCGGTGTGAGAAGTGCACAGAGCCCTGTGCCGA	1020	Qy	653	-----	653
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	Db	2101	ACACCTAGCGAGCGATGCCCAACAGCGGAGATGCGGATCTCTGAAAGAGACGAGAGCTG	2160
Db	1021	GTGTGCTATGCTGTGGGATGGAGCACTTGCAGAGGTGAGGGCAGTTACAGTGCCTAAT	1080	Qy	653	-----	653
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380	Db	2161	AGGAAGTGAAGTGTCTGGGCTTTGGGATCTGGGATCTTACAAGGGCATCTGGATC	2220
Db	1081	ATCCAGAGTTTGTGCTGCAAGAAATCTTTGGGAGCCTTGGCATTTCTGCCGGAGAGC	1140	Qy	653	-----	653
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400	Db	2221	CCTGATCGGGAGAAATGTGAAAATTCAGTGCCCATCAAAGTGTGAGGAAACACATCC	2280
Db	1141	TTTGATGGGACCCAGCCTCCACACTGCCCGCTCCAGCCAGAGAGCTCCAAAGTGT	1200	Qy	653	-----	653
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420	Db	2281	CCCAAGCCAAAGAAATCTTAGAGGAAGCATACGTGATGGCTGTGGCTCCCCA	2340
Db	1201	GAGACTCGGAAGAGATCACAGGTACCTATACATCTCAGCATGGCGGACAGCTGCCT	1260	Qy	653	-----	653
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440	Db	2341	TATGCTCTCCGCTTCTGGGATCTGCTGACATCCAGGTCAGCTGGTGACACAGCTT	2400
Db	1261	GACCTCAGCGTCTCCAGAACCTGCAAGTAATCCGGGAGCAATCTGCAATGCGGCC	1320	Qy	653	-----	653
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460	Db	2401	ATGCCCTATGGCTGCCCTTAGACCATGTCCGGGAAACCCGGGACGCTGGGCTCCAG	2460
Db	1321	TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTGCCCTCAGTGGGAA	1380	Qy	653	-----	653
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480	Db	2461	GACCTGCTGAACGTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGGG	2520
Db	1381	CTGGGCAGTGGACTGGCCCTCATCCACATAACACCCACCTCTGCTTCTGTCGACACGGTG	1440	Qy	653	-----	653
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500	Db	2521	CTCGTACACAGGAGCTTGGCCGCTCGGAACGTGCTGTTCAAGAGTCCCAACCATGTCAAA	2580
Db	1441	CCCTGGGACCAAGCTCTTTCCGAACCCCGCACCAAGCTCTGTCTCCACACTGCGCAACCGGCCA	1500	Qy	653	-----	653
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520	Db	2581	ATTACACACTTCGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGAGTACCATTGCAGAT	2640
Db	1501	GAGGACAGTGTGTGGCGAGGGCTGGCCCTGCCACAGCTGTGGCCCGGAGGCACTGC	1560	Qy	653	-----	653
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540	Db	2641	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCCGCGGCTTTCACC	2700
Db	1561	TGGGTTCCAGGGCCCAACCCAGTGTCAACTGCAGCCAGTTCTTTCGGGGCCAGGAGTGC	1620	Qy	653	-----	653
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560	Db	2701	CACCAGAGTGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2760
Db	1621	GTGGAGGAATGCCAGTACTGCAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	1680	Qy	653	-----	653
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580	Db	2761	AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGTCTGGAAAAAGGGGAGCGG	2820
Db	1681	TTGCCGTGCCACCTGAGTGTGAGCCCAAGTGGCTCAGTGACCTGTTTGGACCGGAG	1740	Qy	653	-----	653
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600				



Dd	61	GCAGGACCCCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCGCAGTCCCGAG	120
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlyAsnLeu	60
Dd	121	ACCCACCTGGACATGCTCCGCCACCTCTACCAAGGCTGCCAGGTGGTGCAGGGAACCTG	180
Qy	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Dd	181	GAACCTACCTTACCTGCCCAACCAATGCCAGCTGTCTTCTGCAAGGATATCCAGGAGGTG	240
Qy	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Dd	241	CAGGCTACGTGCTATCCCTCACAAACAAAGTGAGGAGGTCCCACTGCAGAGGCTGCGG	300
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Dd	301	ATTGTGCGAGGCCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGA	360
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Dd	361	GACCCGCTGAACAATAACCACTGCTACAGGGGCTCCCAAGGAGGCTGCGGGAGCTG	420
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160
Dd	421	CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTCAATCCAGCGGAACCCCGAG	480
Qy	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
Dd	481	CTCTGCTACGAGACAGATTTGTGGAAGGACATCTTCCACAGAACAACCACTGGCT	540
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Dd	541	CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGAAG	600
Qy	201	GlySerArgCysTrpGlyGluSerGluAspCysGlnSerLeuThrArgThrValCys	220
Dd	601	GGCTCCCGCTGCTGGGGAGAGATTCTGAGGATTGTCAGAGCCTGACGCGCACTGTCTGT	660
Qy	221	AlaGlyLysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Dd	661	GCCGTTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGACACGTGT	720
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Dd	721	GCTCGCGGCTGCACGGCCCCAAGCACTCTGACTGCCTGGCTGCCCTCCACTTCAACCA	780
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Dd	781	AGTGGCATCTGTGAGCTGCACCTGCCACGCCCTGGTCACTTCAACACACAGACACATTTGAG	840
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Dd	841	TCCATGCCCAATCCCGAGGGCCGTTATACATTCGGCGCCAGCTGTGTGACTGCTGTGCC	900
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Dd	901	TACAACCTACCTTCTACGACAGTGGGATCTCTACCCCTCGTCTGCCCTGCAACAACAA	960
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Dd	961	GAGGTGACAGCAGAGATGGAAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1020
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Dd	1021	GTGTGCTATGGCTGGCGATGGAGCACTTGGAGAGGTGAGGGCAGTTTACCAGTGGCCAA	1080
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Dd	1081	ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGAGCCTTGGCATTTCTGCCGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Dd	1141	TTTGATGGGACCCAGCTTCAACACTGCGCGCTCCAGCCAGCAGACAGCTCCAAGTGT	1200

```

QY 653 ----- 653
Db 2281 CCCAAAGCCAAACAAAGAAATCTTAGAGAAAGCATACGTGATGCGTGGTGGCTCCCCA 2340
QY 653 ----- 653
Db 2341 TATGTCTCCCGCTTCTGGGCATCTGCTGCATCCACGGTGCAGCTGGTGACACAGCTT 2400
QY 653 ----- 653
Db 2401 ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAACCGCGGAGCCCTGGGCTCCAG 2460
QY 653 ----- 653
Db 2461 GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGGG 2520
QY 653 ----- 653
Db 2521 CTGCTACACAGGAGCTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAAA 2580
QY 653 ----- 653
Db 2581 ATTACAGACTTCGGGCTGCTGGCTGCTGGACATTGACGACACAGATACCATGCAGAT 2640
QY 653 ----- 653
Db 2641 GGGGCAAGTGTCCTCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGCTTCACC 2700
QY 653 ----- 653
Db 2701 CACCAGAGTGATGTGTGGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2760
QY 653 ----- 653
Db 2761 AAACCTTACATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGGAGCGG 2820
QY 653 ----- 653
Db 2821 CTGCCCCAGCCCCCATCTGCACCAATTGATGTCTACATGATCATGTCAAATGTTGGATG 2880
QY 653 ----- 653
Db 2881 ATTGACTCTGAATGTCGCGCAAGATTCCGGGAGTTGGTGCTGTAATTCTCCCGCATGGCC 2940
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 2941 AGGGACCCCGCGCTTTGTGGTCAATCCAGATGAGACTTGGGCCCGCCAGTCCCTTG 3000
QY 665 AspSerThrPheTyrArgSerLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGCACCTTCTPACCGCTCACTGTCTGGAGGACGATGACATGGGGACCTGGTGATGCT 3060
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATCTGGTACCCAGCAGGCGCTTCTTGTCCAGACCTGCCCGCGCGCTGGG 3120
QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
Db 3121 GGATGGTCCACCACAGGACCCGACGCTCATCTACCAGGAGTGGCGTGGGACCTGCACA 3180
QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3181 CTAGGGCTGGAGCCCTCTAGAGGAGGAGCCCGCCAGGTCTCCACTGCACCCCTCCGAGGG 3240
QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db 3241 GCTGGCTCCGATGATTGATGTCACCTGGGAATGGGGGACGCCAAGGGGCTGCAAGC 3300
QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerClnuAspProThrValProLeu 784
Db 3301 CTCCCCACACATAGCCCCAGCCCTCTACAGCGGTACAGTGAGGACCCACAGATCCCCCG 3360
QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804

```

```

Db 3361 CCCTCTGAGACTGATGGCTAGCTTGGCCCTGACCTGCAGCCCGCCAGCTGAATATGTG 3420
QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3421 AACCAAGCCAGATGTTTGGCCCGCCAGCCCTTGGCCCGGAGAGGGCCCTCTGCCTGTGCC 3480
QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3481 CGACCTGCTGGTGGCCACTCTGGAAAGGCCCAAGACTCTCCCCAGGGAGAAATGGGGTC 3540
QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3541 GTCAAAGACGTTTTTCCCTTTGGGGTGCCTGGAGAACCCCGAGTACTTGACACCCAG 3600
QY 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3601 GGAGGAGCTGCCCTCAGCCCGCCACCTCTCTCTGCTTCCCTTCCAGCCAGCTTCGACAACTC 3660
QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3661 TATTACTGGACGAGACCCACCCAGAGCGGGGGCTCCACCCAGACCTTCAAGAGGACA 3720
QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db 3721 CCTACGGCAGAAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG 3765

RESULT 5
ABA92250
ID ABA92250 standard; cDNA; 3768 bp.
XX
AC ABA92250;
XX
DT 17-JUN-2002 (first entry)
XX
DE Human Her-2/neu cDNA.
XX
KW Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
KW receptor; human; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT FT
FT FT
XX
PN W0200212341-A2.
XX
PD 14-FEB-2002.
XX
PF 03-AUG-2001; 2001WO-US24283.
XX
PR 03-AUG-2000; 2000US-0632507.
XX
PA (CORI-) CORIXA CORP.
PA (SMIX ) SMITHKLINE BEECHAM BIOLOGICALS.
PI Cheever MA, Gheysen D;
XX
XX WPI; 2002-241743/29.
XX P-PSDB; AAM51143.
XX
PT Her-2/neu fusion protein for treating or preventing cancer by eliciting
PT or enhancing an immune response to the protein, has Her-2/neu
PT extracellular domain fused to Her-2/neu intracellular or
PT phosphorylation domain -
XX
PS Disclosure; Fig 15; 141pp; English.
XX
CC The present sequence is that of human Her-2/neu oncogene cDNA.
CC The cDNA encodes Her-2/neu (p185), an oncogenic self protein and
CC target for anti-cancer vaccines. The Her-2/neu gene is amplified

```

CC and p185 is overexpressed in a variety of cancers, including breast,  
 CC ovarian, colon, lung and prostate cancer. Her-2/neu is a member  
 CC of the tyrosine kinase family of receptor-like glycoproteins.  
 CC Its overexpression correlates with a poor prognosis in breast and  
 CC ovarian cancers. The invention provides Her-2/neu fusion  
 CC proteins, nucleic acids encoding them, viral vectors, and vaccines  
 CC comprising the fusion proteins or nucleic acid molecules. In  
 CC preferred fusion proteins, the extracellular domain of a Her-2/neu  
 CC protein is fused to a Her-2/neu intracellular domain or  
 CC phosphorylation domain (or its Deltapp fragment). An immune  
 CC response to Her-2/neu protein is elicited or enhanced by  
 CC administering the fusion protein in the form of a vaccine, or by  
 CC transfecting cells of an animal *ex vivo* with a nucleic acid  
 CC encoding the fusion protein, and delivering the transfected cells  
 CC to the animal. The fusion proteins, nucleic acids, and isolated  
 CC specific T-cells are useful for inhibiting the development of a  
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer  
 CC in a patient. T cells that specifically react with a Her-2/neu  
 CC fusion protein can be used to remove tumour cells from a sample in  
 CC order to inhibit the development of cancer in a patient.  
 XX  
 SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

## Alignment Scores:

Pred. No.: 6.12e-227 Length: 3768  
 Score: 4900.00 Matches: 919  
 Percent Similarity: 73.23% Conservative: 0  
 Best Local Similarity: 73.23% Mismatches: 0  
 Query Match: 96.49% Indels: 336  
 DB: 24 Gaps: 1

SE06 (1-919) x ABA92250 (1-3768)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20  
 DB 1 ATGGAGCTGGCGGCTTGGCGCTGGGGGCTCTCTCCGCGCTTGGCCCCGGAGCC 60  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 DB 61 GCGAGCACCAAGTGTGCACCGGCACACATGAAGCTGCGCTCCCTGCCAGCCGAG 120  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrglnGlnCysGlnValValGlnGlyAsnLeu 60  
 DB 121 ACCCACTGGACATGCTCCGCGACCTCTTACCAGGGCTGCCAGGTGGTGCAGGGAAACCTG 180  
 QY 61 GluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80  
 DB 181 GAACCTACCTACCTGCCCCCACCACATGCCAGCTGTCTTCTTCCAGGATATCCAGAGGTG 240  
 QY 81 GlnGlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 DB 241 CAGGCTACGTGCTCATCGCTCACAACTGAGGAGGAGTCCCACTGCAGAGCTGGG 300  
 QY 101 IleValArgGlyThrGlnLeuPheGluaspAsnTyrrAlaLeuAlaValLeuAspAsnGly 120  
 DB 301 ATTGTGGAGGACCCAGCTCTTTGAGGACAACATATGCCCTGGCGTGTAGACAAATGGA 360  
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 DB 361 GACCCGCTGAACATATACCCCTGTACAGGGGCTCCCGAGGAGGCTCGGGAGCTG 420  
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160  
 DB 421 CAGCTTGAAGCCCTCACAGATCTTGAAGAGGGGTCTTGATCCAGCGAACCCCGAG 480  
 QY 161 LeuCysTyrrGlnAspThrIleLeuTrpLysaspIlePheHisLysAsnAsnGlnLeuAla 180  
 DB 481 CTCTGCTACCAAGACACGATTTTGTGAAGACATCTTCCACAAGAACCAACAGCTGGCT 540  
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 DB 541 CTCACATGATAGACCAACCGCTCTGGGGCTTGCACCCCTTCTTCCGATGTGTAAG 600

QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 DB 601 GGCTCCCGCTCTGGGGAGAGAGTTCTGAGGATTGTTCAGAGCCTGACGCGCAGCTCTCTGT 660  
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
 DB 661 GCCGGTGGCTGTGCCCGCTGCAAGGGGCACTGCCCACTGACTGTGCTGCAGCAGTGT 720  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 DB 721 GCTGCCGCTGCACGGGCCCAAGCACTCTGACTCCCTGGCTGCTCCACTTCAACCAAC 780  
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThrPheGlu 280  
 DB 781 AGTGCATCTGTGAGCTGCACTGCCAGCCCTGGTGCACCTACAACACAGACACAGCTTGAG 840  
 QY 281 SerMetProAsnProGluGlyArgTyrrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 DB 841 TCCATGCCCAATCCCGAGGGCGGTATACATTCGGCGCAGCTGTGTGACTGCTGCTGCC 900  
 QY 301 TyrAsnTyrrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 DB 901 TACAACCTACCTTCTACGGAGCTGGGATCCCTGCACCCCTCGTCTGCCCTTGCACAAAC 960  
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 DB 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAGAGTGCAGCAAGCCCTGTGCCGA 1020  
 QY 341 ValCysTyrrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 DB 1021 GTGTGCTATGCTTGGCATGGAGCATTGGCAGAGGTGAGGGCAGTTTACCAGTGCCTAAT 1080  
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 DB 1081 ATCCAGGAGTTTGTCTGGCTGCAAGAAGATCTTGGAGCGCTGGCAATTCTTGGCGGAGAGC 1140  
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
 DB 1141 TTTGATGGGACCCAGCCCTCCACACACTGCCCGCTCCAGCCAGACAGCAGCTCCAAGTGT 1200  
 QY 401 GluThrLeuGluGluIleThrGlyTyrrLeuTyrrIleSerAlaTrpProAspSerLeuPro 420  
 DB 1201 GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGTCGCGCAGACAGCTGCT 1260  
 QY 421 AspleuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 DB 1261 GACCTCAGCTCTTCCAGAACCTCAAGTAATCGGGAGCAATTTCTGCACAAATGGCGCC 1320  
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 DB 1321 TACTCGCTGACCTTCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGAA 1380  
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480  
 DB 1381 CTGGGCACTGAGCTGGCCCTCATCACCATTAACACCCACCTCTGCTTGTGTCACACAGTG 1440  
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuHisThrAlaAsnArgPro 500  
 DB 1441 CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCACCCGGCCA 1500  
 QY 501 GluAspGluCysValGlyGluAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 DB 1501 GAGGACGAGTGTGTGGCGAGGGCTGGCCCTGCCACAGCTGTGCGCCCGAGGGCACTGC 1560  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 DB 1561 TGGGCTCAGGGGCCACCCAGTGTCAACTGCACTGCAGCCAGTTCTTCTCGGGCGCAGAGTGC 1620  
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrrValAsnAlaArgHisCys 560  
 DB 1621 GTGGAGGAATGCCAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGCACTGT 1680  
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580





[illegible]









Db	655	CTCTGCTACAGGACACGATTGTTGGAAAGGACATCTCCACAAGAACACACGCTGGCT	714
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	715	CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG	774
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	775	GGCTCCCGTGTGTGGGAGAGATTCTGAGGATTGTCAGAGCTGACGCGCACTGTCGT	834
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	835	GCGGTGGTGTGCGCGCTGCAAGGGGCCACTGCCCACTGACTGTGCTGCCATGAGCAGTGT	894
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	895	GCTGCCGGTGCACGGGCCCCCAAGCACTCTGACTGCTGGCTGGCTCCCACTCAACCA	954
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	955	AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACCTACAACACACACACGTTTGAG	1014
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	1015	TCAATGCCCAATCCCGAGGGCCGGTATACATTGCGGCCGAGCTGTGTGACTGCTGTCCC	1074
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	1075	TACAACCTACCTTTCTACGGACGTGGGATCCTGCACCCCTGCTGCCCCCTGCACACCAA	1134
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	1135	GAGGTGACAGCAGAGGATGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1194
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1195	GTGTGCTATGGTGTGGGCATGGAGCACCTTGCAGAGGTGAGGCGCATTTACCAAGTGC	1254
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1255	ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGAGAGC	1314
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe	400
Db	1315	TTTGATGGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGCAGCAGCTCCAAGTGT	1374
QY	401	GluThrLeuGluLeuIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1375	GAGACTCTGGAAGATACAGATTACCTATACATCTCAGCATGCGCGCAGACGCTGCCT	1434
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1435	GACCTCAGCGTCTCCAGAACCTGCAAGTAATCCGGGAGCAATTTGACAAATGGCGCC	1494
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1495	TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTGGCTCACTGAGGGAA	1554
QY	461	LeuGlySerClyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1555	CTGGGCAGTGGACTGGCCCTCATCCACCATACACCCACTCTGCTTCGTGCACGCGTG	1614
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1615	CCCTGGGACCACTCTTTCCGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1674
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1675	GAGGACAGTGTGTGGGCGAGGCGCTGCCCTGCCACCACTGTGCCCGGAGGCACTGC	1734
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1735	TGGGGTCCAGGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC	1794
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1795	GTGGAGGAATCCGAGTACTGCAGGGGCTCCCCAGGAGTATGTGAATGCCAGGCACTGT	1854
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1855	TTGCCGTGCCACCCCTGAGTGTGAGCCCAAGATGGTCACTGACCTGTTTGGACCGGAG	1914
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1915	GCTGACCACTGTGTGGCCTGTGCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGC	1974
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1975	CCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAAGTTTCCAGATGAGGAG	2034
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2035	GGCGCATGCGAGCCTTGCCCATCAACTGCACCCACTCCTGTGTGACCTGGATGATGAA	2094
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2095	GGTGTGCCCGCCAGCAGAGAGCCCTCTGACGTCCATCATCTCTCGGTGGTGGC	2154
QY	653	-----	653
Db	2155	ATTCTGCTGCTGCTGTTGGGGTGTCTTTGGGATCCTCATCAAGCAGCGCAGCAG	2214
QY	653	-----	653
Db	2215	AAGATCCGGAAGTACAGATGCGGAGACTGCTGCAGAAACGAGTGTGTGGAGCGCTG	2274
QY	653	-----	653
Db	2275	ACACCTAGCGGAGCGATGCCCAACAGCGCAGATCGGATCCTGAAAGACGACGAGCTG	2334
QY	653	-----	653
Db	2335	AGGAAGTGAAGTGTGTGATCTGGCTGCGCTTTTGGCACAGCTTACAAGGGCATCTGATC	2394
QY	653	-----	653
Db	2395	CCTGATGGGGAGAAATGTGAAAATTCAGTGGCCATCAAAAGTGTGTGAGGAAACACATCC	2454
QY	653	-----	653
Db	2455	CCCAAGCCCAACAAAGAAATCTTAGACGAAGCATACGTATGGCTGTGTGGCTCCCCA	2514
QY	653	-----	653
Db	2515	TATGTCTCCGCTTCTGGGCATCTGCCCTGCATCCACGTCGACGTGGTGACACAGCTT	2574
QY	653	-----	653
Db	2575	ATGCCCTATGCTGCTCTTAGACCATGTCCCGGAAACCGCGGACGCTGGGCTCCCCAG	2634
QY	653	-----	653
Db	2635	GACCTGCTGAAGTGGTGTATGCAGATTGCCAAGGGGATGAGTACCTGGAGGATGTGCGG	2694
QY	653	-----	653
Db	2695	CTCGTACACAGGACTTGGCCGCTCGGAACGCTGCTGGTCAAGAGTCCCACCATGTCAAA	2754
QY	653	-----	653
Db	2755	ATTACAGACTTGGGGCTGGCTGGCTGTGGACATTGACGAGACAGATACCATGCAGAT	2814
QY	653	-----	653
Db	2815	GGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCCTCGCGCGCGGTTACC	2874



Db	355	GAACTCACCCTACCTGCCCCACCAATGCCAGCCCTGCTCTCCTCAGGATATCCAGGAGTG	1435	GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCCGGGAGCAATTCTGCACAATGCGCC
QY	81	GinGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTyrTrpLeuGlyLeuArgSerLeuArgGlu
Db	415	CAGGGCTACGTGCTCATCGCTCACACCAAGTGAGGCAGGTCCACTGCAGAGGTGCGG	1495	TACTCGCTCACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCAGTACGTAGGGAA
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal
Db	475	ATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCGCCTGCTAGACAATGGA	1555	CTGGGCAGTGGACTGGCCCTCATCCACCATAAACCCACCTCTGCTTCTGTCACACGGTG
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro
Db	535	GACCCGCTGAACAATACCACCCCTGTCACAGGGGCTCCCCAGGAGCCCTGGCGAGCTG	1615	CCCTGGGACCACTCTTTTCGGAACCCGACCAAGCTCTGTCACACTGCCAACCCGCCCA
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys
Db	595	CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGTCTTATCAGCGGGAACCCCCAG	1675	GAGGAGAGTGTGTGGCGAGGGCTTGGCTGCCACCAGCTGTGCCCGGAGGGGCACTGC
QY	161	LeuCysTyrGluAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys
Db	655	CTCTGTACAGACACAGATTTTGTGGAAGGACATCTTCCACAAGAACAACAGCTGGCT	1735	TGGGGTCCAGGGCCCAACCCAGTGTCTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTG
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys
Db	715	CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGAAG	1795	GTGGAGGAATGCCAGTACTGCAGGGGCTCCCCAGGAGTATGTGAATGCCAGGCATGT
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu
Db	775	GGCTCCCGTGTGGGGAGAGAGTCTGAGGATGTFCAGAGCTGACGGCGACTGTCTGT	1855	TTGCCGTGCCACCTGAGTGTACGCCCAAGAATGGCTCAGTGACCTGTTTGGACCCGAG
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys
Db	835	GCGGTGGGTGTGCCCGTGCAGGGCCACTGCCACTGACTGTGCTGCATGAGCAGTGT	1915	GCTGACCAGTGTGTGCCCTGTGCCACTATAAGGACCCCTCCCTTCTGCTGGGCCGCTGC
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu
Db	895	GTCGCGGTGCACGGGCCCAAGCACTGTGACTGCTGCGCTGCCCTCCACTCAACCCAC	1975	CCCAGCGTGTGAAACCTGACCTCTCTACATGCCCATCTCGAAGTTTTCAGATGAGGAG
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys
Db	955	AGTGGCATCTGTGAGCTGCACGCCCGCTGGTCACCTACCAACACAGACAGTTTGAG	2035	GGCGCATGCCAGCTTGCCCCCTCACTGCCACCTCTCTGTGTGGACCTTGGATGACAAG
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer
Db	1015	TCCATGCCCAATCCCGAGGCGGTATACATTGCGCGCCAGCTGTGTGACTGCCGTGCC	2095	GGCTGCCCGCGCAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGTGTGTGGC
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	653	-----
Db	1075	TACAACTACCTTTCTACGGACGTGGGATCCTGCACCCCTGCTGCCCTGCACACCA	2155	ATTCTGCTGGTCTGTGCTTTGGGGGTGCTTTTGGGATCCTCATCAAGCGAAGGCAGAG
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	653	-----
Db	1135	GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAAGGCCCTGTGCCCA	2215	AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGAAACGGAGCTGTGGAGCCGCTG
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	653	-----
Db	1195	GTGTGTATGTCTGGGCATGGAGCACTTGCAGAGAGTGAGGGCAGTTACCAAGTGCCAAT	2275	ACACCTAGCGGAGCGATGCCCAACAGCGCAGATCGGGATCCTGAAAGAGACGGAGCTG
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	653	-----
Db	1255	ATCCAGAGTGTGTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTCGCGGAGAGC	2335	AGGAAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	653	-----
Db	1315	TTTGATGGGGACCCAGCCTCCAACTGCTCCCGCTCCAGCCAGAGCAGCTCCAAAGTGT	2395	CCTGATGGGGAGATGTGAAAATTCAGTGGCCATCAAAGTGTGTAGGGAACAACATCC
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	653	-----
Db	1375	GAGACTCTGGAAGAGATCAGGTTACCTATACATCTCAGCATGGCGGACAGCCTGCCT	2455	CCCAAAGCCAACAAGAAATCTTAGACGAAGCATACGTGCTGCTGTGGGCTCCCCA
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	653	-----
			2515	TATGCTCCCCCTTCTGGGCATCTGCGCTGACATCCACGGTGACGTGTGACACAGCTT

QY 653 ----- 653  
 Db 2575 ATGCCCTATGGCTGCCCTTAGACCATGTCCGGGAAACCGCGGACGCTGGGGTCCAG 2634  
 QY 653 ----- 653  
 Db 2635 GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGTACCTGGAGGATGTGCGG 2694  
 QY 653 ----- 653  
 Db 2695 CTCGTACACAGGGACTTGGCGCTCGGAACGTGCTGTCAAGAGTCCCAACCATGTCAAA 2754  
 QY 653 ----- 653  
 Db 2755 ATTACAGACTTCGGGCTGGCTCGGCTCTGGACATTCACGACAGAGTACCATGCAGAT 2814  
 QY 653 ----- 653  
 Db 2815 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTTCTCCGCGCGGTTCACC 2874  
 QY 653 ----- 653  
 Db 2875 CACCAGAGTGATGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2934  
 QY 653 ----- 653  
 Db 2935 AAACCTTACGATGGATCCAGCCCGGGAGATCCCTGACCTGTGGAAAGGGGACGG 2994  
 QY 653 ----- 653  
 Db 2995 CTGCCCCAGCCCCCATCTGCACCATGTGATGTCTACATGATCATGTGTCAAATGTGGATG 3054  
 QY 653 ----- 653  
 Db 3055 ATTGACTCTGAATGTCGGCCAAGATTCGGGAGTTGGTGTCTGAATTTCTCCCGCATGGCC 3114  
 QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664  
 Db 3115 AGGAGCCCCAGCGCTTTGGTTCATCCAGATGAGACTTGGCCCGCAGCTCCCTTG 3174  
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684  
 Db 3175 GACAGCACCTTCTACCGCTCCTGCTGGAGACGATGACATGGGGACCTGGTGGATGCT 3234  
 QY 685 GluLeuTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704  
 Db 3235 GAGAGTATCTGGTACCCCGACAGCGGGCTTCTGTCTCCAGACCTGCCCGCGCGCTGGG 3294  
 QY 705 GlyMetValHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724  
 Db 3295 GGCATGCTCCACACAGCCAGCCGCTCATCTACACAGGAGTGGCGTGGGGACCTGACA 3354  
 QY 725 LeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGluGly 744  
 Db 3355 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCGCCCGCAGGTCTCCACTGGCACCTCCGAAGGG 3414  
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764  
 Db 3415 GCTGGCTCCCATGATGATTTGATGTGACCTGGGAATGGGGACGCCAAGGGGCTGCAAGC 3474  
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
 Db 3475 CTCCCCACACATGACCCAGCCCTCTACACGGGTACAGTGAGGACCCACAGTACCCCTG 3534  
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
 Db 3535 CCCTCTGAGACTGATGCTAGTTGCCCCCTGACCTGCAGCCCCCAGCCTGATATGTG 3594  
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824  
 Db 3595 AACCAGCCAGATGTGGCCCCAGCCCTTCGCCCGGAGAGGGGCTCTGCTGCTGCTGCC 3654

825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
 Db 3655 CGACCTGCTGGTCCACTCTGGAAGGCCCAAGACTCTCTCCCGAGGAGATGGGGTC 3714  
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
 Db 3715 GTCAAGACGCTTTTGGCTTTGGGGTGGCGTGAGAACCCCGAGTACTTGCACACCCAG 3774  
 QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884  
 Db 3775 GGAGGAGCTCCCTCAGCCCCACCTCTCTGCTTCCAGCCAGCTTCGACAACTC 3834  
 QY 885 TyrTyrTyrAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
 Db 3835 TATTACTGGACACGAGGCCACACAGAGCGGGGGCTCCACCAGCACCTTCAAAGGACA 3894  
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
 Db 3895 CCTACGCAGAGAACCAGAGTACCTGGGTCTGGACGTGCCAGTG 3939  
 RESULT 10  
 ABZ34969  
 ID ABZ34969 standard; cDNA; 4473 BP.  
 XX AC ABZ34969;  
 XX DT 05-FEB-2003 (first entry)  
 XX XX Human gene expression profile polynucleotide SEQ ID NO 81.  
 DE Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;  
 KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;  
 KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;  
 KW gene expression; gene; ss.  
 XX OS Homo sapiens.  
 XX WO200274979-A2.  
 XX PD 26-SEP-2002.  
 XX PF 20-MAR-2002; 2002WO-US08456.  
 XX PR 20-MAR-2001; 2001US-276947P.  
 XX PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.  
 XX PI Wan J, Wang Y;  
 XX WI; 2002-740862/80.  
 XX New gene expression profile generated from primary, endothelial,  
 PT epithelial, and muscle cell types, useful for identifying disease  
 PT pathologies involving alterations of gene expression, e.g. cancer  
 XX  
 PS Claim 3; Page 274-276; 850pp; English.  
 XX  
 CC The invention relates to a gene expression profile comprising one or more  
 CC genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type  
 CC is a coronary artery endothelium, umbilical artery or vein endothelium,  
 CC aortic endothelium, dermal microvascular endothelium, pulmonary artery  
 CC endothelium, myometrium microvascular endothelium, keratinocyte  
 CC epithelium, bronchial epithelium, mammary epithelium, prostate  
 CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,  
 CC small airway epithelium, renal epithelium, umbilical artery smooth  
 CC muscle, neonatal dermal fibroblast, pulmonary artery smooth  
 CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,  
 CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,  
 CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,  
 CC osteoblasts or prostate stromal cell. The gene expression profile is used  
 CC for determining the level of RNA expression for a sample, determining the  
 CC phenotype of a cell and distinguishing cell types. The gene or a protein  
 CC expression profile is useful in identifying disease pathologies

CC involving alterations of gene expression. The assessment of expression  
 CC profiles may provide meaningful information with respect to tumour type  
 CC and stage, treatment methods, and prognosis. The gene or protein  
 CC expression profile may also be used for creating microarrays. The  
 CC microarray is useful for genetic and physical mapping of genomes, DNA  
 CC sequencing, genetic or medical diagnosis, genotyping of organisms,  
 CC confirming cell or tissue identifications and in identifying promising  
 CC antibiotics, antiviral or antifungal agents.

XX  
 SQ Sequence 4473 BP; 902 A; 1383 C; 1329 G; 859 T; 0 other;

Alignment Scores:  
 Pred. No.: 7,31e-227 Length: 4473  
 Score: 4900.00 Matches: 919  
 Percent Similarity: 73.23% Conservative: 0  
 Best Local Similarity: 73.23% Mismatches: 0  
 Query Match: 96.49% Indels: 336  
 DB: 24 Gaps: 1

SE06 (1-919) x AB234969 (1-4473)

QY 1 MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20  
 DB 175 ATGGAGCTGGCGGCGCTTGGCGCGCTGGGGGCTCTCTCGCCCTCTTGGCCCGCGAGCC 234  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 DB 235 GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 294  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
 DB 295 ACCCACTGGACATGCTCGCCACCTCTACCAAGGCTGCCAGGTGGTGGAGGAACCTG 354  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
 DB 355 GAACCTACCTACTGCTCCACCACTGCGGCTGCTCTTCCCTCAGGATATCCAGAGGTG 414  
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 DB 415 CAGGGCTACGTCTCATGCTCACAACCAAGTGAAGGAGGCTCCCACTGCAGAGGCTGCG 474  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 DB 475 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCTGCGGCTGCTAGACAATGGA 534  
 QY 121 AspProLeuAsnAspThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 DB 535 GACCCGCTGAACAATACACCCCTGTACAGGGGCTCCCAAGAGGCTGCGGGAGCTG 594  
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160  
 DB 595 CAGCTTCGAAGCTCACAGAGATCTTGAAGGAGGGTCTTGATCCAGCGGAACCCCGAG 654  
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180  
 DB 655 CTTCTGCTACCGAGACAGATTTTGTGAAGGACATCTTCCACAAGAACAACAGCTGGCT 714  
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 DB 715 CTCACACTGATAGACACACACCCGCTCTCTGGGCTGCCACCCCTGTTCTCCGATGTGA 774  
 QY 201 GlySerArgCysTrpGlyLeuSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 DB 775 GGCTCCCGCTGCTGGGAGAGATTCTGAGGATTCTCAGAGGCTGACGCGCTGCTCTGT 834  
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCys 240  
 DB 835 GCGGTGGCTGTGTCGCTGCAAGGGGCGCACTGCCCACTGACTGCTGCTGCTGCTGCTG 894  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 DB 895 GCTGCGGCTGCACGGGCCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTG 954

QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 DB 955 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACACACACACACAGCTTTGAG 1014  
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 DB 1015 TCCATGCCCAATCCCGAGGCGGTATACATTCCGGCGCAGCTGTGTACTGCTGCTGCC 1074  
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 DB 1075 TACAACCTACCTTCTACGAGACGTGGATCTGCACCTCTGCTGCTGCTGCTGCTGCTG 1134  
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 DB 1135 GAGGTGCACAGCAGAGATGGAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCGA 1194  
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 DB 1195 CTGTGCTATGTGCTGTGGCATGGAGCACTTGCAGAGAGTGAAGGCGAGTTACCAAGTGC 1254  
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 DB 1255 ATCCAGAGATTGTGCTGCTGCAAGAGATCTTTGGAGCCTGGCATTTGTCGGGAGAGC 1314  
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
 DB 1315 TTTGATGGGAGCCAGCCTCCAAACACTGCCCGCTCCAGCCAGCAGCAGCTCCCAAGTGT 1374  
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420  
 DB 1375 GAGACTCTGGAAGAGATCACAGTTTACCTATATCATCTCAGCATGCGCGGACAGCTGCT 1434  
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 DB 1435 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGACAGTAATCTGCAATGCGGCC 1494  
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 DB 1495 TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGTGGGCTGCGCTCACTGAGGAA 1554  
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480  
 DB 1555 CTGGCAGCTGGACTGGCCCTCATCCACCATAACACCCCTCTGCTGCTGCTGCTGCTG 1614  
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 DB 1615 CCCTGGGACCAGCTCTTTCGGAAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCA 1674  
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 DB 1675 GAGGACGAGTGTGTGGGCGAGGCTGCGCTGCCACCACTGCTGCGCGCGAGGCACTGC 1734  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 DB 1735 TGGGTTCAGGCGCCACCCAGTGTGTCAACTGCACCCAGTTCCTTTCGGGGCCAGAGTGC 1794  
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
 DB 1795 GTGGAGGAATGCCAGTACTGCGAGGGCTCCCGGAGGAGTATGTAAGTCCAGGCACTGT 1854  
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 DB 1855 TTGCGGTGCCACCTGAGTGTGAGCCCAAGATGCTCAGTACCTGTTTGGACCGGAG 1914  
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
 DB 1915 GCTGACCAAGTGTGTGGCTGTGCCACTATAAGGACCTTCCCTTCTGCTGCTGCTGCTG 1974  
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
 DB 1975 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCACTCTGGAAGTTTCCAGATGAGGAG 2034  
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640







QY 501 GluaspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaAraGlyHisCys 520  
 Db 1675 GAGGACGAGTGTGGGGAGGGCTGGCTGCCACCAAGCTGTGGCCGAGGCACTGC 1734  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnCys 540  
 Db 1735 TGGGCTCAGGGCCCAACCACTGTCTAACTGCAGCCAGTTCTTCGGGGCCAGAGTGC 1794  
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaAraHisCys 560  
 Db 1795 GTGGAGGAATGCCGAGTACTCAGGGCTCCCGAGGAGTATGTGAATGCCAGCACTGT 1854  
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 Db 1855 TTGGCGTCCCACTTGTAGTGTGACCCAGCAATGGCTCAGTCACTGTTTGGACCGGAG 1914  
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaAraCys 600  
 Db 1915 GCTGACCAAGTGTGGGCTGTGCCCACTATAGGACCTCTCTCTGGTGGCCGCTGC 1974  
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
 Db 1975 CCCAGCGGTGAAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2034  
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640  
 Db 2035 GGCATGCCAGCTTGGCCCATCAACTGCACCCACTCTCTGTGGACCTGGATGACAAG 2094  
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653  
 Db 2095 GGCTGCCCGCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGGCGTGGTTGGC 2154  
 QY 653 ----- 653  
 Db 2155 ATTCTGCTGCTGGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGACGAG 2214  
 QY 653 ----- 653  
 Db 2215 AAGATCCGGAAAGTACAGGATCGGAGACTGCTCGAGGAAGGAGTGGTGGAGCCGCTG 2274  
 QY 653 ----- 653  
 Db 2275 ACACCTAGCGAGCGATGCCCAACAGCGCAGATGCGGATCTGAAAGACGAGAGCTG 2334  
 QY 653 ----- 653  
 Db 2335 AGGAAGTGAAGTGTGGATCTGGCTTTTGGCACAGTCTACAAAGGATCTGGATC 2394  
 QY 653 ----- 653  
 Db 2395 CCTGATGGGAGAAATGTGAAATTTCCAGTGGCCATCAAAAGTGTGAGGGAAACACATCC 2454  
 QY 653 ----- 653  
 Db 2455 CCCAAGCCAAAGAANAATTTAGACGAAGCATACTGATGTGGTGTGGGTCCCCA 2514  
 QY 653 ----- 653  
 Db 2515 TATGCTCCCGCTTCTGGGATCTGCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2574  
 QY 653 ----- 653  
 Db 2575 ATGCCCTATGGCTGCCCTTAGACCATGTCCGGGAAACCGGACGCCCTGGGCTCCAG 2634  
 QY 653 ----- 653  
 Db 2635 GACCTGTGAAGTGTATGACATTGCCAAGGGGATGAGCTACCTGGAGGATGTGGG 2694  
 QY 653 ----- 653  
 Db 2695 CTCGTACACAGGACTTGGCGCTCGGACGTGTGGTCAAGAGTCCCAACCATGTCAA 2754  
 QY 653 ----- 653

Db 2755 ATTACAGACTTTCGGCTGGCTGGCTGGACATTGACGACAGAGTACCATGCAGAT 2814  
 QY 653 ----- 653  
 Db 2815 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGGGGGTTAC 2874  
 QY 653 ----- 653  
 Db 2875 CACCAGAGTGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2934  
 QY 653 ----- 653  
 Db 2935 AAACCTTACGATGGGATCCAGCCGGGAGATCCCTGACCTGCTGAAAAAGGGGAGCG 2994  
 QY 653 ----- 653  
 Db 2995 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGGTCAAAATGTTGGATG 3054  
 QY 653 ----- 653  
 Db 3055 ATTGACTCTGAATGTGCGCCCAAGATTCCGGGAGTTGGTGTCTGAATTTCTCCGCATGGCC 3114  
 QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664  
 Db 3115 AGGACCCCCAGCGCTTTGTGTCATCCAGAAATGAGGACTTGGGCCACGACGTCCTTG 3174  
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684  
 Db 3175 GACAGCACTTCTACCGCTCACTGCTGGAGACCATGACATGGGGGACCTGGTGGATGCT 3234  
 QY 685 GluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGlyAlaGly 704  
 Db 3235 GAGGAGTATCTGGTACCCACAGCGCTTCTTCTGTCAGACCTGCCCCGGCGCTGGG 3294  
 QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724  
 Db 3295 GGCATGTGTCCACCAACAGCAGCCAGCTCATCTACAGAGTGGCGGTGGGACCTGACA 3354  
 QY 725 LeuGluGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744  
 Db 3355 CTAGGGCTGGAGCCCTCTGAAGAGAGGGCCCCAGGTCTCCACTGGCACCTCCGANGG 3414  
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764  
 Db 3415 GCTGGCTCCGATGTATTGTGATGTGACCTGGGAATGGGGAGCAGCAAGGGCTGCAAGC 3474  
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
 Db 3475 CTCCCCACATGACCCAGCCCTCTACAGCGGTACAGTAGAGAGCCACACAGTACCCCTG 3534  
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
 Db 3535 CCCTCTGAGACTGATGGCTACGTTGGCCCCCTGACCTGACGCGCCCGACCCCTGAAATGTG 3594  
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824  
 Db 3595 AACCAGCCAGATTTCCGCCCCAGCCCTTCGCCCCGAGAGGGCCCTCTGCCCTGCTGCC 3654  
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
 Db 3655 CGACCTGTGTGGTCACTGTGAAAGGCCCAAGACTCTCTCCCGAGGAGAAATGGGTC 3714  
 QY 845 ValLysAspValPheAlaPheGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
 Db 3715 GTCAAGACGTTTGTGCTTTGGGGTGGCTGTGAGAACCCCGAGTACTTGTACACCCAG 3774  
 QY 865 GlyGlyAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884  
 Db 3775 GGAGAGCTGCCCTCAGCCCCACCTCTCTGCTTCCAGCCAGCCTTCGACACCTC 3834  
 QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
 Db 904

Db 3835 TATTACTGGACAGACCCACAGAGCGGGGGCTCCACCAGCACCTTCAAAGGGACA 3894

QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
 |||||  
 Db 3895 CCTACGGCAGAAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG 3939

RESULT 12  
 AAA09455  
 ID AAA09455 standard; DNA; 3768 BP.  
 XX  
 AC AAA09455;  
 XX  
 DT 10-AUG-2000 (first entry)  
 XX  
 DE Human heregulin 2 (Her2) coding sequence.  
 XX  
 KW Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;  
 KW self-protein; cell-associated peptide antigen; foreign epitope;  
 KW cancer; breast cancer; prostate cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200020027-A2.  
 XX  
 PD 13-APR-2000.  
 XX  
 PF 05-OCT-1999; 99WO-DK00525.  
 XX  
 PR 05-OCT-1998; 98DK-0001261.  
 PR 20-OCT-1998; 98US-0105011.  
 XX  
 PA (MEBI-) M & E BIOTECH AS.  
 XX  
 PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;  
 PI Gautam A, Birk P, Karlsson G;  
 XX  
 DR WPI; 2000-349917/30.  
 DR P-PSDB; AAY92620.  
 XX  
 PT Inducing immune responses to weakly immunogenic, tumor associated  
 PT peptide antigens for the treatment of breast and prostate cancer  
 XX  
 PS Claim 62; Page 187-193; 220pp; English.  
 XX  
 CC The claims detail a method for inducing immune responses against weakly  
 CC immunogenic cell-associated peptide antigens (PA) such as those  
 CC associated with cancers (i.e. self-proteins), for example, human  
 CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or  
 CC fibroblast growth factor 8b (FGF8b). The method comprises effecting  
 CC simultaneous presentation by antigen producing cells (APCs) of the  
 CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)  
 CC group derived from the PA and/or at least 1 B-cell group derived from the  
 CC cell-associated PA; and (2) at least 1 first T helper cell group which is  
 CC foreign to the animal. Analogues of human PSM, human Her2 and  
 CC human/murine FGF8b comprising a substantial part of all known and  
 CC predicted CTL and B-cell epitopes of the respective PA and including at  
 CC least one foreign T helper epitope are also claimed. The method is used  
 CC to treat prostate, prostate/breast or breast cancer when the PA is human  
 CC FSM, FGF8b and Her2, respectively.  
 XX  
 SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:  
 Pred. No.: 1,48e-226 Length: 3768  
 Score: 4892.00 Matches: 918  
 Percent Similarity: 73.15% Conservative: 0  
 Best Local Similarity: 73.15% Mismatches: 1  
 Query Match: 96.34% Indels: 336  
 DB: 21 Gaps: 1

SE06 (1-919) x AAA09455 (1-3768)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuProGlyAla 20  
 |||||

Db  
 1 ATGGAGCTGGCGCCTTGTGCGCTGGGGCTCCTCTCGCCTCTTGCCTCCCGGAGCC 60  
 |||||  
 QY  
 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 |||||  
 Db  
 61 GCAGCACCACCAAGTGTGACCGGCACAGACATGAAGCTCGGCTCCCTCCAGTCCCGAG 120  
 |||||  
 QY  
 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
 |||||  
 Db  
 121 ACCCAGCTGGACATGCTCGCCACCTCTTACCAAGGGCTGCCAGGTGGTGGAGGAAACCTG 180  
 |||||  
 QY  
 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80  
 |||||  
 Db  
 181 GAACTCACCTACCTGCCACCAATGCCAGCTGTCTCTCTCCAGGATATCCAGGAGGTG 240  
 |||||  
 QY  
 81 GlnGlyTyrValLeuIleAlaHisGlnValArgGlnValProLeuGlnArgLeuArg 100  
 |||||  
 Db  
 241 CAGGGCTAGTGTCTATCGCTCACAAACCAAGTGAGCAGGTCCACGTGAGAGGCTGCGG 300  
 |||||  
 QY  
 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 |||||  
 Db  
 301 ATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGTACAAATGGA 360  
 |||||  
 QY  
 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 |||||  
 Db  
 361 GACCCGCTGAACAATACCACCTGTGTACAGGGGCTCCCCAGGAGGCTCGGGAGGTG 420  
 |||||  
 QY  
 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgsnProGln 160  
 |||||  
 Db  
 421 CAGCTTCGAAGCTTCACAGAGATCTTGAAGAGGGGTCTTGTATCCAGCGGAACCCCCAG 480  
 |||||  
 QY  
 161 LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeuAla 180  
 |||||  
 Db  
 481 CTCCTGCTACCAGACACAGATTTTGTGAAGAGGACATCTCCACAAGAACACACAGCTGGCT 540  
 |||||  
 QY  
 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 |||||  
 Db  
 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCGATGTGTAAG 600  
 |||||  
 QY  
 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 |||||  
 Db  
 601 GGCTCCCGCTGCTGGGGAGAGAGATTCTGAGGATTGTACAGAGCTGACGGCGCACTGTCTGT 660  
 |||||  
 QY  
 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
 |||||  
 Db  
 661 GCCGTGGCTGTGCCGCTGCAAGGGCCACCTGCCACCTGCTGCTGCCATGAGCAGTGT 720  
 |||||  
 QY  
 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 |||||  
 Db  
 721 GCTGCCGGCTGCACGGCCCCCAAGCACTCTGACTGCTGCCCTGCCCTCCACTTCAACCCAC 780  
 |||||  
 QY  
 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 |||||  
 Db  
 781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACAAACACAGACACAGCTTTGAG 840  
 |||||  
 QY  
 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 |||||  
 Db  
 841 TCCATGCCCAATCCCGAGGGCCGGTATACATTGGGGCCAGCTGTGTGACTGCCTGTCCC 900  
 |||||  
 QY  
 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 |||||  
 Db  
 901 TACAACCTACCTTTCTACGGACGCTGGGATCTCTGCACCTCGCTGCCCTCCGACACCA 960  
 |||||  
 QY  
 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 |||||  
 Db  
 961 GAGGTGCACAGCAGAGGATGAAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCGCCGA 1020  
 |||||  
 QY  
 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 |||||  
 Db  
 1021 GTGTGCTATGTGCTGGGCATGGAGCACTTGCAGAGAGGTGAGGCGAGTTACAGTGCCTAAT 1080  
 |||||  
 QY  
 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 |||||

Db	1081	ATCCAGAGTTTGTGCTGCGTCCAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe	400
Db	1141	TTTTGATGGGACCAGCCTCCAACACTGCCCGCTCCAGCCAGAGACAGCTCCAAGTGT	1200
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGACAGCTGCCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCTCTTCCAGAACCTGCAAGTAATCCGGGAGCAATTCGCACAATGGCGCC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTGGCGCTCACTGAGGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGCAGTGGACTGGCGCTCATCCACCATAAACACCACCTCTGCTTCGTGCACACGGTG	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCAGCTCTTTTCGGAACCCCGCACCAAGCTCTGTCCACACTGCCAACCGGCCA	1500
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACAGTGTGGGCGAGGGCTGGCTGCCACCAAGCTGTGCGGCCCGAGGGCACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGCCACCAGTGTCAACTGCAGCCAGTTCTTCTGGGGCCAGGAGTGC	1620
Qy	541	ValGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGAAATGCCAGTACTGCAAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACGTG	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCTGAGTGTACGCCCCAGCAATGGCTCAGTGACCTGTTTTGGACCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCAGTGTGGCGCTGTGCCACTATAAGSACCCTCCCTCTCGTGGCGCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCCAGCGGTGTGAACGTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGCATGCCAGCTTGCCCCATCACTGCACCCACCTCTCTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCCGACGAGAGAGCCCTCTCAGCTCCATCTGCTCGCGGTGGTGGC	1980
Qy	653	-----	653
Db	1981	ATTCTGCTGCTGCTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGCGACAG	2040
Qy	653	-----	653
Db	2041	AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAACGGAGCTGGTGGAGCCGCTG	2100
Qy	653	-----	653
Db	2101	ACACCTAGCGGCGGATGCCCAACAGCGCAGATGCGGATCCTGAAAGAGAGCGGAGCTG	2160
Qy	653	-----	653
Db	2161	AGGAAGTGAAGTGTGGATCTGGCGCTTTTGGCACAGTCTTACAAGGGCATCTGGATC	2220
Qy	653	-----	653
Db	2221	CCTGATGGGAGAGATGTGAAAAATTCAGTGGCCATCAAAAGTGTTCAGGGAAAAACACATCC	2280
Qy	653	-----	653
Db	2281	CCCAAGCCAAACAAAGAAATCTTAGAGGAAGCATACGTGATGCTGTGTGGGCTCCCCA	2340
Qy	653	-----	653
Db	2341	TATGCTCTCCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2400
Qy	653	-----	653
Db	2401	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCCGCGACGCTGGGCTCCAG	2460
Qy	653	-----	653
Db	2461	GACCTGCTGAATGGTGTATGTCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG	2520
Qy	653	-----	653
Db	2521	CTCGTACACAGGAGACTTGGCGCTCGGAACGTGCTGTCAAGAGTCCCAACCATGTCAAA	2580
Qy	653	-----	653
Db	2581	ATTACAGACTTGGGGCTGGCTCGGCTGCTGGACATTTCACGACAGACAGATACCATGCCAGAT	2640
Qy	653	-----	653
Db	2641	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGCTTCACC	2700
Qy	653	-----	653
Db	2701	CACCAGAGTATGTGTGGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2760
Qy	653	-----	653
Db	2761	AAACCTTACGATGGATCCCGCGGAGATCCCTGACCTGTCTGGAAGGGGGAGCGG	2820
Qy	653	-----	653
Db	2821	CTGCCCCAGCCCCCATCTGCACCAATTGATGCTTACATGATCATGTGTCAAATGTTGGATG	2880
Qy	653	-----	653
Db	2881	ATTGACTCTGAATGTGCGGCCAAGATTCCGGGAGTTGTGTGTAATTCCTCCCGCATGGCC	2940
Qy	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	2941	AGGACCCCCCAGCGCTTTGTGGTCAATCCAGAATGAGGACTTGGCCCCAGCAGTCCCTTG	3000
Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACCTTCTACCGCTCAGTGTGGAGACGATGACATGGGGACCTGGTGGATGCT	3060
Qy	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3061	GAGAGTATCTGGTACCCAGCAGGGCTTCTTGTCCAGACCTTCCCCGGGCGCTGGG	3120
Qy	705	GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr	724
Db	3121	GGCATGGTCCACCACAGCCAGCCAGCTCATCTACCAAGAGTGGCGTGGGACCTGCACA	3180
Qy	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
Db	3181	CTAGGCGTGGAGCCCTCTGAAGAGGAGGCCCCAGGTCTCCACTGGCACCTCCGAGGG	3240
Qy	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer	764
Db	3241	GCTGGCTCCGATGATTTTGTGTGACCTGGGAATGGGGGACGCCAAGGGGCTGCAAGC	3300

QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
 Db 3301 CTCCTGAGACTGATGGCTACCTGACAGCGGTACAGTGGAGAGCCACACAGTACCCCTG 3360  
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
 Db 3361 CCCTCTGAGACTGATGGCTACCTGACAGCGGTACAGTGGAGAGCCACACAGTACCCCTG 3420  
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824  
 Db 3421 AACCGACCGAGTTCGGCCCGAGCCCTTCGCCCGAGAGGGCCCTCTGCTGCTGCC 3480  
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
 Db 3481 CGACCTGCTGGTGCACCTCTGGAAGGGCCAAAGACTCTCTCCCGAGGGAAGTGGGTC 3540  
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
 Db 3541 GTCAAGACGTTTTCCTTGGGCTGGCGTGGAGAACCCCGAGTACTTGCACCCCGAG 3600  
 QY 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884  
 Db 3601 GGAGGAGCTGCCCTCAGCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3660  
 QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
 Db 3661 TATTACTGGACAGGACCCACAGAGCGGGGCTCCACCCAGCACCTTCAAGGGACA 3720  
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
 Db 3721 CCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGAGTCCGACGTG 3765  
 RESULT 13  
 ABZ35744  
 ID ABZ35744 standard; DNA; 3768 BP.  
 XX  
 AC ABZ35744;  
 XX  
 DT 07-FEB-2003 (first entry)  
 XX  
 DE Human ERBB2 polynucleotide SEQ ID NO 52.  
 XX  
 KW Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;  
 KW protozoas; gene expression; antisense; tumour; infection; plasmodium;  
 KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;  
 KW Hepatitis C virus; human papilloma virus; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 DE DE10100588-A1.  
 PN  
 XX  
 PD 18-JUL-2002.  
 XX  
 PF 09-JAN-2001; 2001DE-1000588.  
 PR  
 XX  
 PR 09-JAN-2001; 2001DE-1000588.  
 XX  
 PA (RIBO-) RIBOPHARMA AG.  
 XX  
 PI Kreutzer R, Limmer S, Rost S, Hadwiger P;  
 XX  
 DR WPI; 2002-683450/74.  
 XX  
 DR Inhibiting expression of target genes, useful e.g. for treating tumors,  
 PT by introducing into cells two double-stranded RNAs that are  
 PT complementary to the target  
 XX  
 PS Claim 13; Page 38-39; 100pp; German.  
 XX  
 CC The invention relates to inhibiting expression of a target gene in a cell  
 CC by introducing at least two oligonucleotides (dsRNAi and II), both  
 CC with a double-stranded (ds) structure of at most 49 sequential nucleotide  
 CC pairs. At least part of one strand (S1, S2) of the ds structures in each

CC of dsRNAi and II are complementary to regions in the target gene. The  
 CC method uses antisense inhibition of gene expression using double stranded  
 CC RNA inhibition (RNAi). The method is particularly used to treat tumours  
 CC or infections, especially by plasmodium or viruses/viroids (pathogenic on  
 CC humans, animals or plants). The method provides more effective inhibition  
 CC of expression than known methods using a single dsRNA, even at very low  
 CC concentrations. When dsRNA has at least one unpaired nucleotide at the  
 CC end, stability (and thus effective concentration in the cell) is  
 CC improved and efficiency can be increased further by pretreating the cells  
 CC with interferon. The present sequence is that of a target DNA of the  
 XX invention.

XX Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.48e-226 Length: 3768  
 Score: 4892.00 Matches: 918  
 Percent Similarity: 73.15% Conservative: 0  
 Best Local Similarity: 73.15% Mismatches: 1  
 Query Match: 96.34% Indels: 336  
 DB: 24 Gaps: 1

SEQ6 (1-919) x ABZ35744 (1-3768)

QY 1 MetGluLeuAlaAlaLeuCysArgTyrGlyLeuLeuLeuAlaLeuLeuProGlyAla 20  
 Db 1 ATGGAGCTGGCGCCCTGTGGCGCTGGGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProLysSerProGlu 40  
 Db 61 GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGTCTCGGCTCTCTCTCTCTCTCTCTCT 120  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu 60  
 Db 121 ACCACCTGGAGCATGCTCGCCACCTCTTACCAGGCTGCCAGGTGTGGAGGAACCTG 180  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
 Db 181 GAACCTCACCTACCTGCCACCAATGCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 Db 241 CAGGGCTAGCTCTCATCTGCACCAAGTGAAGGAGGTCCACTGCAGAGGTGCGG 300  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 Db 301 ATTGTGCGAGGACCCAGCTCTTTGAGGACACTATGCCCTGCGCTGTACACATGGA 360  
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 Db 361 GACCCGCTGAACAATACACACCTCTGTACAGGGGCTCTCCAGGAGGCTCGGGAGCTG 420  
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160  
 Db 421 CAGCTTCGAAGCTTCACAGAGATCTTGAAGAGGGGTCTTGTATCAGAGGGAACCCCGAG 480  
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180  
 Db 481 CTCTGCTACCGACACAGATTTTGTGAAGAGACATCTCCACAGAACACACAGCTGGCT 540  
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCTCTGTCTCCGATGTGTAAG 600  
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 Db 601 GGCTCCGCTGCTGGGGAGAGAGTCTGTAGAGATTGTACAGAGCTGACGCGCACTGTCTGT 660  
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240  
 Db 661 GCCGTGGCTGTGCCCGCTGCAGGGGCGACCTGCCACCTGCTGCTGCTGCTGCTGCTGCTG 720  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260



```

QY      654  -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db      2941 AGGGACCCAGCGCTTGTGTCATCCAGAAATAGGACTTGGGCCAGCCAGTCCCTTG 3000
QY      665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db      3001 GACAGACCTCTTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGTTGATGCT 3060
QY      685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db      3061 GAGGAGTATCTGTATCCAGAGGCTTCTGTCTCCAGACCCCTGCCCCGGCGCTGGG 3120
QY      705 GlyMetValHisHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724
Db      3121 GGCATGCTCCACACAGGACCGCAGCTCATCTACGAGGAGTGGCGTGGGACCTGACA 3180
QY      725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db      3181 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCCAGGTCTCCACTGGCACCCCTCCGAAGG 3240
QY      745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
Db      3241 GCTGGCTCGATGATTTGATGTCACCTGGGAATGGGGGACCCAAAGGGCTGCAAGC 3300
QY      765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db      3301 CTCCCCACACATGACCCCGCCTCTACAGCGGTACAGTGAGGACCCACACATACCCCTG 3360
QY      785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db      3361 CCCTCTGAGACTGATGGCTACCTTGCCTCCCTGACCTGCAGCCCCCAGCCTGAATATGTG 3420
QY      805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db      3421 AACCAGCCAGATGTTGGGCCCCAGCCCCCTTCGCCCGAGAGGGCCCTCTGCTGCTGCC 3480
QY      825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db      3481 CGACCTGCTGTTGCCACTCTGAAAGGGGCCAAGACTCTCTCCCGAGGGAAGTGGGTC 3540
QY      845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db      3541 GTCAAAGACGTTTTTGCCCTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACCCCCAG 3600
QY      865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
Db      3601 GGAGGAGCTGCCCTCAGCCCCACCTCTCTCTGCTTCAGCCAGCCCTTCACACACCTC 3660
QY      885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db      3661 TATTACTGGGACAGGACCCACAGAGCGGGGGCTCCACCCAGCACCTTCAAAGGGACA 3720
QY      905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db      3721 CCTACGGCAGAACCCAGAGTACCTGGGTCTGGAGTCCAGTGCCAGTG 3765

RESULT 14
ABX09987
ID ABX09987 standard; DNA; 3768 BP.
XX AC ABX09987;
XX AC ABX09987;
DT 23-JAN-2003 (first entry)
XX DE Human ERBB2 DNA fragment SEQ ID 52.
XX DE Oligoribonucleotide; interferon; oncogene; cytokine; id; developmental;
KW KW prion; inhibition; human; ds.
XX OS Homo sapiens.
XX PN DE10100587-C1.

```

```

XX 21-NOV-2002.
PD 09-JAN-2001; 2001DE-1000587.
XX PF 09-JAN-2001; 2001DE-1000587.
XX PR 09-JAN-2001; 2001DE-1000587.
XX PA (RIBO-) RIBOPHARMA AG.
XX PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX WPI; 2002-742209/81.
DR Inhibiting expression of target genes, e.g. oncogenes, in cells, by
PT introduction of complementary double-stranded oligoribonucleotide,
PT after treating the cell with interferon
PT
XX Disclosure; Page 43-44; 98pp; German.
XX This invention describes a novel method for inhibiting expression of a
CC target gene by introducing into the cell that contains the target gene
CC at least one oligoribonucleotide (dsRNA) that has a double-stranded
CC (ds) structure of not more than 49 consecutive nucleotides (nt), where
CC at least a segment of one strand of the ds structure is complementary
CC with the target gene and the cells are treated with interferon before
CC introduction of dsRNA. The method is used to inhibit expression of
CC target genes, particularly oncogenes, cytokine genes, id (not defined)
CC protein genes; developmental or prion genes, or genes expressed in
CC pathogenic organisms (particularly plasmidia) or in viruses or viroids
CC (pathogenic in humans, animals or plants). Treating the cells with
CC interferon greatly increases the extent to which dsRNA can inhibit
CC expression of the target genes, and the effect is even greater when dsRNA
CC are modified to increase their stability. ABX09936-ABX10075 represent
CC gene fragments used to illustrate the method of the invention.
XX Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;
SQ

```

```

Alignment Scores:
Pred. No.: 1.48e-226 Length: 3768
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservatives: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 24 Gaps: 1
SEQ6 (1-919) x ABX09987 (1-3768)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGGAGCTGGCGGCTTGTGGCGCTGGGGCTCTCTCGCCCTCTTGGCCCCCGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGCACCAAGTGTGCACCGGCACAGATGAGCTGCGGCTCCCTCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCCACCTGGACATGCTCCGCCACCTCTACCAAGGCTGCCAGGTGGTGCAGGAAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 181 GAACTCACCTACTGCCACCAATGCCCTGTCTCTCTCCAGGATATCCAGGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTACGTGCTCATGCTCAACAAGTGAAGGAGGAGTCCCTGAGAGGCTGCCG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGCTGTAGACAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140

```



Db	361	GACCCGCTGAACATACACCCCTGTCTACAGGGGCGCTCCCCAGGAGGCGCTCGCGGAGCTG	420
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160
Db	421	CAGCTTCGAAGCTTCAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCCAG	480
Qy	161	LeuCystyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
Db	481	CTCTGCTACCAGGACAGATTTGTGGAAGGACATCTCCACAAGAACACACAGCTGGCT	540
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	541	CTCACATGATAGACACCAACCCCTCTCGGGCGCTGCCACCCCTGTCTCCGATGTGTAA	600
Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	601	GGCTCCCGCTGTGGGAGAGATTTCTGAGGATTTTCAGAGCTGACGCCCTGACCGCACCTGCTGT	660
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	661	GCCGGTGGCTGTGCCCTGCAAGGGGCCACTGCCACTGACTGCTGCCATGAGCAGTGT	720
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	721	GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCCCTGGCCTGCCCTTCAACCCAC	780
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThrPheGlu	280
Db	781	AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCCACCTACCAACACAGACAGCTTTGAG	840
Qy	281	SerMetProAsnProGluGlyArgTyrrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	841	TCCATGCCCAATCCCGAGGGCCGCTATACATTCGGGCCAGCTCTGTGACTGCTGTGCC	900
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACCTACCTTTCTACGGAGCTGGGATCTCGCACCTCTGCTGCCCTTGCACAAACAA	960
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACAGCAGGAGTGAACACACGGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGGA	1020
Qy	341	ValCystyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGCTGGGCATGGAGCACTTCGGAGAGGTGAGGCGAGTACCAGTGGCAAT	1080
Qy	361	IleGlnIlePheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGGAGTTGCTGGCTGCAAGAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1141	TTTGATGGGAGCCAGCCTCCACACATGCCCGCTCCAGGCAGAGCAGCTCCAAGTGT	1200
Qy	401	GluThrLeuGluIleThrGlyTyrLeuTyrrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTGGAAGAGATCAGAGTTACCTATACATCTACCATGGCGGAGACGCTGCCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCTCTTCCAGAACCCTCAAGTAAATCCGGGAGCAATTTCTGCACAAATGGCGC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCTTGCAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCCTCAGTGAGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGCAGTGGAGCTTCATCCACCATACACCCACTCTGCTTCGTGTCACACGGTG	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCACTCTTTCGGAACCCCGCACCAAGCTCTGTCTCCACACTGCGCAACCGCCA	1500
Qy	501	GluAspGluCysValGlyGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGAGCAGTGTGTGGCGAGGCGCTGGCCCTGCCACACAGCTGTGGCCCGAGGCGACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGCCCAACCCAGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC	1620
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGAAATGCCAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACCTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCCCTGAGTGTAGCCCCAGCAATGGCTCAGTACCTGTTTTGGACCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCACTGTGGCGCTGTGCCACTATAAGGACCCCTCCCTTCTCGTGGCGCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAATTTCCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGCATGCCAGCTTGCCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCCGAGCAGAGAGCAGCCCTCTGACGTCCATCGTCTCGCGGTGGTTGGC	1980
Qy	653	-----	653
Db	1981	ATTCTGCTGCTGCTGTCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGCAGCAG	2040
Qy	653	-----	653
Db	2041	AAGATCGGGAAGTACAGGATGCGGAGACTCTCGCAGAAAGGAGCTGTGGAGCCCGCTG	2100
Qy	653	-----	653
Db	2101	ACACCTAGCGGAGCGATGCCCAACAGCGCGCAGATCGGATCCTCTGAAAGAGAGCGGAGCTG	2160
Qy	653	-----	653
Db	2161	AGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2220
Qy	653	-----	653
Db	2221	CCTGATGGGGAGAATGTGAAAATTCAGTGGCCATCAAACTGTTGAGGGAAACACATCC	2280
Qy	653	-----	653
Db	2281	CCCAAGCCCAACAAAGAAATCTTACAGGAAGCATACGTGATGGCTGGTGTGGGCTCCCA	2340
Qy	653	-----	653
Db	2341	TATGCTCTCCGCGCTTCTGGGCATCTGCTGACATCCACGGTGACAGCTGTGTGACACAGCTT	2400
Qy	653	-----	653
Db	2401	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCGCGGACGCCCTGGGCTCCCAG	2460
Qy	653	-----	653
Db	2461	GACCTGCTGAACGTGTATGCAGATTGCCAAGGGGATGAGCTACCTTGGAGGATGTGGCG	2520
Qy	653	-----	653
Db	2521	CTCGTACACAGGAGCTTGGCGCTCGGAACGCTGTGTGTCAGAGTCCCAACCATGTCAAA	2580

QY 653 ----- 653  
 Db 2581 ATTACAGACTTCGGCTGGCTGGCTGGACATTGACGACAGACAGTACCATGCAGAT 2640  
 QY 653 ----- 653  
 Db 2641 GGGGGCAAGGTGCCATCAAGTGGATGGGCTGGAGTCCATTCTCCGCGGGGGTTCAAC 2700  
 QY 653 ----- 653  
 Db 2701 CACCAGATGATGTGGAGTATGTTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2760  
 QY 653 ----- 653  
 Db 2761 AAACCTTACGATGGGATCCAGCCGGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2820  
 QY 653 ----- 653  
 Db 2821 CTGCCCCAGCCCCCATCTGCACATTGATGTCTACATGATCATGTCAAATGTTGGATG 2880  
 QY 653 ----- 653  
 Db 2881 ATTGACTCTGAATGTCCGCCAAGATTCCGGGAGTTGGTCTCTGAATTCCTCCGCATGGCC 2940  
 QY 654 -----  
 Db 2941 AGGGACCCCGAGCGCTTTGTGGTTCATCCAGAATGAGGACTTGGGCCACCCAGTCCCTTG 3000  
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684  
 Db 3001 GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGGTGGTGT 3060  
 QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704  
 Db 3061 GAGGAGTACTGGTACCCAGCAGCGCTTCTCTGCCAGACCTGCCCGGGCGCTGGG 3120  
 QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724  
 Db 3121 GGCATGGTCCACCACAGCCAGCCAGCTCATCTACCAGGAGTGGCGGTGGGACCTGACA 3180  
 QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744  
 Db 3181 CTAGGCGTGGAGCCCTCTGAAGAGGAGGCCGCCAGGTCTCCACTGGCACCCCTCCGAAGG 3240  
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764  
 Db 3241 GCTGGCTCGAGTATTTGATGGTACCTGGGATGGGGGACCCAGGGGCTGCANAGC 3300  
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
 Db 3301 CTCCCCACACATGACCCCGCCCTCTACAGCGGTACAGTGAGGACCCACAGTACCCCTG 3360  
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
 Db 3361 CCCTCTGAGACTGATGGCTACGTTGCCCTGACCTGACCTGACGCCCCAGCCTGAATATGTG 3420  
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824  
 Db 3421 AACCACCCAGATGTTCCGCCCCAGCCCTTCCGCCCGAGAGGCCCTCTGCTGTCTGCC 3480  
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
 Db 3481 CGACCTGCTGGTCCACTCTGGAAGGGCCAGACTCTCTCCCGAGGAGAAATGGGGTC 3540  
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
 Db 3541 GTCAAGAGCTTTTTCCTTGGGGTGGCTGGAGAACCCCGAGTACATTCACACCCAG 3600  
 QY 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884  
 Db 3601 GGAGGAGCTGCCCTCAGCCCCACCTCTCTCTGCTTCAGCCCTTCAGCCCTTCAGAACCTC 3660  
 QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904

Db 3661 TATTACTGGGACGAGCCACGAGCGGGGCTCCACCCAGCACCTTCAAAGGACA 3720  
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
 Db 3721 CCTACGGCAGAACCCAGAGTACCTGGGTCTGGAGTCCAGTG 3765  
 RESULT 15  
 AAD43935  
 ID AAD43935 standard; cDNA; 3768 BP.  
 AC AAD43935;  
 DT 13-DEC-2002 (first entry)  
 DE Human HER-2 cDNA.  
 XX  
 KW Transgenic animal; transgenic; mammary gland cell; HER2; tumour;  
 KW cancer; therapy; apoptosis; cytostatic; human; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 CDS 1..3768  
 FT /\*tag= a  
 FT /product= "Human HER2 protein"  
 XX  
 PN US2002035736-A1.  
 XX  
 PD 21-MAR-2002.  
 XX  
 PF 16-MAR-2001; 2001US-0811115.  
 PR 16-MAR-2000; 2000US-189844P.  
 XX  
 PA (ERIC/) ERICKSON S.  
 PA (KING/) KING K.  
 PA (SCHW/) SCHWALL R.  
 XX  
 PI Erickson S, King K, Schwall R;  
 XX  
 DR WPI; 2002-401155/43.  
 DR P-PSDB; AAE26349.  
 XX  
 PT New transgenic non-human mammal that produces detectable levels of a  
 PT native human HER2 protein in its mammary gland cells, useful as tumor  
 PT models for testing HER2-directed cancer therapies, and for identifying  
 PT anticancer agents  
 XX  
 PS Example 2; Page 24-26; 83pp; English.  
 XX  
 CC The invention relates to a transgenic non-human mammal that produces in  
 CC its mammary gland cells detectable levels of a native human HER2 protein  
 CC or its fragment. The transgenic animals are useful as tumour models for  
 CC testing HER2-directed cancer therapies, and for identifying anticancer  
 CC agents. The animals may also be used as source of cells which can be  
 CC immortalised in culture, in screening for compounds that have potential  
 CC as prophylactic or therapeutic treatments of diseases or disorders  
 CC involving expression of HER2. The anti-cancer molecules are useful for  
 CC inducing apoptosis or cell death of cancer cells. The present sequence  
 CC is human HER-2 cDNA.  
 XX  
 SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1.48e-226 Length: 3768  
 Score: 4892.00 Matches: 918  
 Percent Similarity: 73.15% Conservative: 0  
 Best Local Similarity: 73.15% Mismatches: 1  
 Query Match: 96.34% Indels: 336  
 DB: 24 Gaps: 1  
 SEQ6 (1-919) x AAD43935 (1-3768)

```
Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyVala 20
Db 1 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCTCTCGCCCTTGTGCCCGGAGCC 60
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGCACCCAAAGTGTGCACCGGCACACATGAAGCTGGCGCTCCCTGCCAGTCCCGAG 120
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCACCTGGACATGCTCCCGCACTCTACAGGGCTGCCAGGTGGTGAGGGAAACCTG 180
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 181 GAACCTACCTACCTGCCACCAATGCCAGCTGTCTCTGTCAGGATATCCAGGAGTG 240
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTACGTGCTCATCGCTCACAACTGAGGAGGCTCCCACTGCAGAGGCTGCGG 300
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATGTGGGAGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGGA 360
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCGTGACATACCAACCCCTGTACAGGGCCCTCCCAAGAGGCTGCGGGAGCTG 420
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTTGAAGCTTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCGAG 480
Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCGTGCTACCAAGGACAGATTTGTGAAGAGGATCTTCCACAAGAACCAACCACTGGCT 540
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 600
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCCTCCGCTGCTGGGAGAGAGTTCAGAGATTCAGAGCTTCAGAGCTTCAGCGGCACTGTCTGT 660
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCCGCTGGCTGTGCCCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 720
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTCGCGCTGCACGGGCCCAACCACTCTGACTGCTGCCCTGCCCTCCACTTCAACCAAC 780
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGGCATCTGTGAGCTGCATGCCAGCCCTGGTCCACTACAAACACAGACACACGTTTGA 840
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCCATGCCCAATCCAGGGCCGGTATACATTCGGCGCCAGCTGTGACTGGCTGTCCC 900
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 TACAACCTACCTTCTACGGAGCTGGATCTCGACCCCTCGTCTGCCCTTCGACAAACCAA 960
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGAAAGCCCTGTGCCCGA 1020
Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1021 GTGTGCTATGTGCTGGCAGTGGAGCACTTGCAGAGGTGAGGGCAGTTACCACTGGCAAT 1080
Qy 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1081 ATCCAGGAGTTGCTGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC 1140
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1141 TTTGATGGGACCCAGCTCCAAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT 1200
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420
Db 1201 GAGACTCTGGAAGAGATCAGAGTTACTATACATCTCAGCATGCCGAGAGCCTGCCT 1260
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1261 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGAGCAATTTCTGCACAATGCGGCC 1320
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1321 TACTCGCTGACCTTGCAGGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGAA 1380
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 CTGGGAGTGGAGTGGCCCTCATCCACATAACACCCACTCTCTCTGTCACACCGGTG 1440
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCCTGGGACCACTCTTTCGGAACCCGACCAACTCTGCTCCACACTGCCAACCGGCCA 1500
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGGACGAGTGTGGGCGAGGGCTGGCCCTGCCACCACTGTGCGCCGAGGCACTGC 1560
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGCTCAGGGCCCACTGAGTGTCACTGACGAGCAGTTCTTCGGGGCCAGGAGTGC 1620
Qy 541 ValGluCysArgValLeuGlnGlyLeuProArgGlyTyrValAsnAlaArgHisCys 560
Db 1621 GTGGAGGAATGCCAGTACTGCGAGGGCTCCCCAGGAGATATGTGAATGCCAGCACTGT 1680
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCGCTGCCACCTGAGTGTGAGCCCAAGTGGCTCAGTGACCTGTTTGGACCGAG 1740
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys 600
Db 1741 GCTGACCACTGTGGGCTGTGCCCACTATAAGACCCCTCCCTTCTGCGTGGCCCGCTGC 1800
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGGTGTGAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1861 GGGCATGTGCCAGCTTGGCCCATCACTGCACCCCACTCTCTGTGTGGAGCTGGATGACAAG 1920
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCCGAGCAGAGAGCCCTCTGACGTCCATCTGCTGCGGTGGTGGC 1980
Qy 653 653 1981 ATCTGCTGCTGCTGGTCTTTGGGGTGGTCTTTGGGATCTCATCAAGCGCAGCAGCAG 2040
Qy 653 653 2041 AAGATCCGGAAGTACAGATCGGAGACTGCTCGAGGAAACGGAGCTGGTGGAGCCGCTG 2100
Qy 653 653 2101 ACACCTAGCGGAGCGATGCCAACCCAGGCGCAGATCGGATCTCTGAAAGACGAGGAGCTG 2160
Qy 653 653 2161 653
```

```

Db      2161 AGGAAGGTGAAGTGCTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2220
QY      653  -----
Db      2221 CCTGATGGGAGAAATGTGAAAATTCAGATGGCCATCAAAAGTGTGAGGAAAACACATCC 2280
QY      653  -----
Db      2281 CCCAAAGCCCAACAAGAAATCTTAGACGAAGCATACGTGATGCTGGTGTGGGCTCCCA 2340
QY      653  -----
Db      2341 TATGTCCTCCGCTTCTTGCGCATCTGCTTGACATCCACGGTGACAGTGTGACAGCTT 2400
QY      653  -----
Db      2401 ATGCCCTATGGCTCTCTTAGACCATGTCCGGGAAAACCGCGGAGCGCTGGGCTCCAG 2460
QY      653  -----
Db      2461 GACCTGCTCAACTGGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGGG 2520
QY      653  -----
Db      2521 CTCGTACACAGGACTTGGCCGCTCGGAACGTGCTGTTCAAGAGTCCCAACCATGTCAA 2580
QY      653  -----
Db      2581 ATTACAGACTTCGGGCTGGCTCGGCTGTGGACATTGACGACAGACAGTACCATGCAGAT 2640
QY      653  -----
Db      2641 GGGGCAAGTGTCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGGCGGCTTCAAC 2700
QY      653  -----
Db      2701 CACCAGAGTGTGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGCGC 2760
QY      653  -----
Db      2761 AAACCTTACATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2820
QY      653  -----
Db      2821 CTGCCCGAGCCCGCATCTGCACCATTTGATGCTTACATGATCATGTGTCAAATGTTGGATG 2880
QY      653  -----
Db      2881 ATTGACTCTGAATGTGGGCCAAGATTCCGGGAGTTGGTCTCTGAATTCCTCCGCGATGGCC 2940
QY      654  -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db      2941 AGGGACCCCGCGCTTTGTGTCTATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTG 3000
QY      665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db      3001 GACAGCACCTTCTACCGCTCACTGTGGAGGACGATGACATGGGGACCTGGTGATGCT 3060
QY      685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db      3061 GAGGAGTATCTGTGTACCCAGAGGGGCTTCTTCTGTCCAGACCTCGCCCGGCGCTGGG 3120
QY      705 GlyMetValHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724
Db      3121 GGCATGGTCCACACAGGACCGCGACGCTCTACAGGAGTGGCGGTGGGACCTGCACA 3180
QY      725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db      3181 CTAGGCTGGAGCCCTCTCAAGAGGAGGCCCCAGGCTCTCCACTGGCACCTCCGAAGGG 3240
QY      745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764

```

```

Db      3241 CCTGGCTCCGATGTATTGATGGTGACCTGGGAATGGGGCAGCCAAAGGGGCTGCAAGC 3300
QY      765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
Db      3301 CTCCCCACACATGACCCCGAGCCCTCTACAGCGGTACAGTAGGACCCACACATACCCCTG 3360
QY      785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db      3361 CCCTCTGAGACTGAGTGTGCTGCTGCCCCCTGACCTGCAGCCCCCAGCCTGAATATGG 3420
QY      805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db      3421 AACCAGCCAGATGTTGCGCCCGAGCCCTTCCGCCCGAGAGGGCCCTTGCTGCTGCC 3480
QY      825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db      3481 CGACCTGCTGTGGTCCACTCTGGAAGGGCCCAAGACTCTCTCCCAGGGGAAGATGGGTC 3540
QY      845 ValLysAspValPheAlaPheGlyGlyAlaValAlaGluAsnProGluTyrLeuThrProGln 864
Db      3541 GTCAAAGACGTTTTTGGCTTTGGGGTGCGGTGGAGAACCCGAGTACTTGACACCCAG 3600
QY      865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884
Db      3601 GGAGGAGTGGCCCTCAAGCCACCTCTCTCTGCTTTCAGCCCGAGCCTTCACACACCTC 3660
QY      885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db      3661 TATTACTGGACGAGGACCCAGAGCGGGGGGCTCCACCCAGCACCTTCAAGGGGACA 3720
QY      905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
Db      3721 CCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGAGCTGGCCAGTG 3765

```

Search completed: October 15, 2003, 23:52:57  
 Job time : 1038.87 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:09:13 ; Search time 13234.4 Seconds  
(without alignments)  
2840.777 Million cell updates/sec

Title: SEQ6  
Perfect score: 5078  
Sequence: 1 MELAALCRWGLLLALLPPGA.....TFRGTPTAENPEYLGLDVPV 919

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool/P/HOLLERAN480/runat\_15102003\_131912\_20536/app\_query.fasta\_1.4685  
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZB=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 -CGN\_1\_1\_22724-runat\_15102003\_131912\_20536 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NGB\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	4900	96.5	3768	6	AR034479	AR034479 Sequence
2	4900	96.5	3768	6	AX201817	AX201817 Sequence
3	4900	96.5	3768	6	AX380923	AX380923 Sequence
4	4900	96.5	3768	6	AX384604	AX384604 Sequence
5	4900	96.5	3768	6	AX465456	AX465456 Sequence
6	4900	96.5	4473	6	AR080259	AR080259 Sequence
7	4900	96.5	4473	6	AR167390	AR167390 Sequence
8	4900	96.5	4473	9	HSEB2R	X03363 Human c-erb
9	4892	96.3	3768	6	AX060704	AX060704 Sequence
10	4892	96.3	3768	6	AX467229	AX467229 Sequence
11	4892	96.3	3768	6	AX481438	AX481438 Sequence
12	4892	96.3	4530	6	AR202597	AR202597 Sequence
13	4892	96.3	4530	6	AR283481	AR283481 Sequence
14	4892	96.3	4530	6	AX282577	AX282577 Sequence
15	4892	96.3	4530	6	AX587649	AX587649 Sequence
16	4892	96.3	4530	6	AX644071	AX644071 Sequence
17	4892	96.3	4530	6	BD005474	BD005474 Cellular
18	4892	96.3	4530	6	121124	121124 Sequence 9
19	4892	96.3	4530	6	159745	159745 Sequence 9
20	4892	96.3	4530	9	HUMHER2A	M11730 Human tyros
21	4892	96.3	9274	6	AX060703	AX060703 Sequence
22	4722	93.0	3678	6	AX505114	AX505114 Sequence
23	4438	87.4	3780	4	AB008451	AB008451 Canis fam
24	4309.5	84.9	2763	6	AX380942	AX380942 Sequence
25	4309.5	84.9	2781	6	AX380944	AX380944 Sequence
26	4142.5	81.6	4062	10	HAMNEU	D16295 Mesocricetu
27	4138.5	81.5	4727	10	AY116182	AY116182 Rattus no
28	4134	81.4	3955	6	AX380924	AX380924 Sequence
29	4134	81.4	3955	6	121129	121129 Sequence 14
30	4134	81.4	3955	6	159750	159750 Sequence 14
31	4134	81.4	3955	10	RNNEUR	X03362 Rat mRNA fo
32	4125.5	81.2	3771	6	AX189662	AX189662 Sequence
33	4125.5	81.2	3771	6	AX380925	AX380925 Sequence
34	4112	81.0	4694	10	BC046811	BC046811 Mus muscu
35	4112	81.0	4695	10	BC053078	BC053078 Mus muscu
36	3632	71.5	2385	6	AR082744	AR082744 Sequence
37	3632	71.5	2385	6	AR099963	AR099963 Sequence
38	3632	71.5	2385	6	AR143949	AR143949 Sequence
39	3525	69.4	2871	6	AX023363	AX023363 Sequence
40	3471	68.4	1872	6	AR181000	AR181000 Sequence
41	2585	50.9	1692	6	AX268287	AX268287 Sequence
42	2585	50.9	2091	6	AX268288	AX268288 Sequence
43	2573.5	50.7	1665	6	AX268285	AX268285 Sequence
44	2573.5	50.7	2070	6	AX268286	AX268286 Sequence
45	1874.5	36.9	1316	9	AF177761	AF177761 Homo sapi

ALIGNMENTS

RESULT 1

AR034479	AR034479	Sequence 1 from patent US 5869445.	3768 bp	DNA	linear	PAT 29-SEP-1999
LOCUS	Sequence 1	from patent US 5869445.	3768 bp	DNA	linear	PAT 29-SEP-1999
DEFINITION	Sequence 1	from patent US 5869445.	3768 bp	DNA	linear	PAT 29-SEP-1999
ACCESSION	AR034479					
VERSION	AR034479.1	GI:5950084				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 3768)					
AUTHORS	Cheever,M.A. and Disis,M.L.					
TITLE	Methods for eliciting or enhancing reactivity to HER-2/neu protein					
JOURNAL	Patent: US 5869445-A 1 09-FEB-1999;					
FEATURES	Location/Qualifiers					
source	1..3768					
BASE COUNT	759 a 1171 c 1119 g 719 t					
ORIGIN	/organism="unknown"					
Alignment Scores:						
Pred. No.:	5.39e-187					
Score:	4900.00					
Percent Similarity:	73.23%					
Best Local Similarity:	73.23%					
Query Match:	96.49%					
DB:	6					
SEQ6 (1-919) x AR034479 (1-3768)						
Qy	1	MetcluleuAlaAlaLeuCysArqTrpGlyLeuLeuLeuAlaLeuProGlyAla	20			
Db	1	ATGAGCTGGCGCCTTGCCGCTGGGGCTCTCTCGCCCTTTGCCCGCGGAGCC	60			
Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40			
Db	61	GCGAGCACCAAGTGTGCACGGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG	120			
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrglnClyCysGlnValValGlnGlyAsnLeu	60			
Db	121	ACCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTCAGGGAACCTG	180			
Qy	61	GluLeuThrTyrlLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80			
Db	181	GAACCTACCTACCTGCCACCAATGCCAGCCTGTCTCTTCGCAGGATATCCAGAGGTG	240			
Qy	81	GlnGlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100			
Db	241	CAGGGCTACGTGCTCATCGCTCAACCAAGTCAGGAGGTCCCACTGCAGAGGCTCGCG	300			
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAspAsnGly	120			
Db	301	ATTGTGGAGGACCCAGCTCTTTGAGGACAACATATGCCCTGGCCGTGTAGACAATGGA	360			
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140			
Db	361	GACCCGTGACATATACCACCCCTGTACAGGGGCCCTCCCAGAGAGCCCTGCGGAGCTG	420			
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLeuGlyGlyValLeuIleGlnArgAsnProGln	160			
Db	421	CAGCTTCGAAGCCCTCACAGAGATCTTGAAGAGAGGGTCTTGATCCAGCGGAACCCAC	480			
Qy	161	LeuCystyrGlnAspThrIleLeuTrpIlyAspIlePheHisLysAsnAsnGlnLeuAla	180			
Db	481	CTCTGCTTACCAGGACAGATTTTGTGAAGAGACATCTTCCACAAGAACACACCTGGCT	540			
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200			
Db	541	CTCACACTGATAGACCAACCGCTCTCTGGGCCCTGCCACCCCTGTTCTCCGATGCTGA	600			
Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220			
Db	601	GGCTCCCGCTGCTGGGAGAGAGTCTGAGGATGTCTAGAGCCCTGACGCGCAGCTGTCTGT	660			







```
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCCTGGGACCAAGCTCTTTGGGAACCCGACCAAGCTCTGCTCCACACTGCGCAACCGGCCA 1500
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGACAGAGTGTGGGGAGGGCTGGCCCTGCCACCAAGCTGTGCGGCCGAGGCACTGC 1560
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGGTCCAGGGCCACCCAGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGAGTGC 1620
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGlyLeuValAsnAlaArgHisCys 560
Db 1621 GTGGAGGAATGCCAGTACTCAGGGGCTCCCGAGGAGTATGTGAATGCCAGCACTGT 1680
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCCGTGCCACCTGAGTGTACGCCCCAGAAATGGCTCAGTGACCTGTTTTGGACCGGAG 1740
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrIlysAspProPheCysValAlaArgCys 600
Db 1741 GCTGACAGTGTGGGCTGTGCCCACTATAGGACCCCTCCCTTCTCGGTGGCCCGCTGC 1800
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG 1860
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 1861 GGCCCATGCCAGCTTCCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 1920
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCGAGCAGAGAGCCGCCCTCTGACGTCCTATCTCTGCGGTGGTTGGC 1980
Qy 653 2000
Db 1981 ATTCGTGTCGTGTCGTTGGGGTGTCTTTGGGATCCTCATCAAGCGACGACGAG 2040
Qy 653 2100
Db 2041 AAGATCCGGAGTACACGATGCGGAGACTGCTGCGAAGAACGGAGCTGGTGGACCGCTG 2100
Qy 653 2200
Db 2101 ACACCTAGCGAGCGATGCCCAACCGGCGAGATCGGGATCCTGAAAGACAGGAGCTG 2160
Qy 653 2220
Db 2161 AGGAAGGTGAAGTGTGGATCTGGCCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2220
Qy 653 2280
Db 2221 CCTGATGGGAGAAATGTGAATAATCCAGTGGCCATCAAAAGTGTGAGGGAAAAACACATCC 2280
Qy 653 2340
Db 2281 CCCAAACCAACAAAGAAATCTTAGACGAAGCATACGTGATGGCTGTGTGGGCTCCCCA 2340
Qy 653 2400
Db 2341 TATGTCCTCCGCCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACACTT 2400
Qy 653 2460
Db 2401 ATGCCCTATGGCTGCCCTTAGACCATGTCGGGAAAAACCGCGAGCCCTGGGCTCCAG 2460
Qy 653 2520
Db 2461 GACCTGCTGAAGTGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG 2520
Qy 653 2580
Db 2581 ATTACAGACTTCGGGCTGGCTCGCTGCTGAGCATTTGACGAGACAGAGTACCATGCAGAT 2640
Qy 653 2700
Db 2701 CACCAGAGTGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGSCC 2760
Qy 653 2820
Db 2821 CTGCCCCAGCCCCCATCTGCACCATTTGATGCTTACATGATCATGGTCAAAATGTTGGATG 2880
Qy 653 2940
Db 2941 AGGACCCCCAGCCCTTTTGGTCAATCAGAAATGAGGACTTTGGGCCCGCAGCAGTCCCTTG 3000
Qy 653 3060
Db 3061 GAGGAGTATCTGGTACCCAGCAGCGGCTTCTTCTCCAGACCTGCCCCGGCGCTGGG 3120
Qy 653 3180
Db 3181 GGCATGTCTCCACACAGCGCCAGCTCTACAGAGGAGTGGCGGTGGGAGCTTGACA 3240
Qy 653 3300
Db 3301 CTAGGGCTGGAGCCCTCTGAAGAGAGGCCCCCAGGCTCTCCACTGGCACCTCCGAAGGG 3360
Qy 653 3420
Db 3421 AACCAGCCAGATGTTGGCCCCAGCCCCCTTCGCCCGAGAGGGCCCTCTGCTGCTGCC 3480
Qy 653 3540
Db 3541 GTCAAGACGCTTTTGGCTTTGGGGTGGCTGGAGAACCCCGAGTACTTTGACACCCAG 3600
Qy 653 3660
Db 3661 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAsnLeu 3720
```





```

QY 725 LeuGlyLeuProSerGluGluGluAlaProArgSerProLeuAlaProSerGluGly 744
|||||
Db 3181 CTAGGCGTGGAGCCCTCTGAGAGAGAGGCCCCAGGTCTCCACTGCACCCCTCCGAGGG 3240

QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764
|||||
Db 3241 GCTGGCTCCGATGATTGATGCTGACCTGGGAATGGGGACGCCAAGGGCTGCAAGC 3300

QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784
|||||
Db 3301 CTCCTCACACATACCCCGCCCTCTACAGCGGTACAGTGAGAGACCCACAGTACCCCTG 3360

QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
|||||
Db 3361 CCTCTGAGACTGATGGCTAGCTGGCCCTGACCTGCAGCCCCAGCCTGATATG 3420

QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
|||||
Db 3421 AACGAGCCAGATGTCGGCCCGCCAGCCCTTCGCCCGGAGAGGGCCCTCTGCTGCTGCC 3480

QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
|||||
Db 3481 CGACCTGCTGGTCCCTCTGAAAGAGGCCCAAGACTCTCTCCCGAGGAGAAATGGGTC 3540

QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
|||||
Db 3541 GTCAAAAGACGTTTTCCTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCAG 3600

QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884
|||||
Db 3601 GGAGGAGCTGCCCTCAGCCCGCCACCTCTCTCTGCTTCCAGCCAGCCTTGACAAACCTC 3660

QY 885 TyrTyrTriAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904
|||||
Db 3661 TATTACTGGACGAGACCCAGACAGCGGGGGCTCCACCCAGCACCTTCAAGGGGACA 3720

QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
|||||
Db 3721 CCTACGCGAGAGAACCAGAGTACCTGGGTCTGGAGTGCCAGTG 3765

RESULT 4
AX384604
LOCUS
DEFINITION
Sequence 1 from Patent WO0214503.
ACCESSION
AX384604
VERSION
AX384604.1 GI:19577806
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Hand-zimmermann,S., Cheever,M.A., Foy,T.M., Lodes,M.J., Kalos,M.D.,
McNeill,P.D. and Vedvick,T.S.
TITLE
Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL
Patent: WO 0214503-A 1 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
location/Qualifiers
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:19577807"
/translation="MELAALCRWGLLLPAGAASVCTGTGDKMLRLPASPETHLD
MRHLVQGVQVQGNLELYLPTNASLSLFDIQEVOGVLYLAHQVROVPLQRLRV
RGQLFEDNIALVDLNGDPLNNTTPTVTCASPGGRELQLRLSTELKKGGLIQNRNQ
LCYQDTILMKDIFHKNNQLALTIDTNRSRACHPCSPMCKGSRWCGESSEDCQSLTRT
VCAGGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNT
DTFSPNPEGRVTFGASCVTAPYNYLSTDVSGCTLCVPLHNOEVTABDGTORCEKC
KESARVCYGLMEHLREYRANTSANIOFAGCKTFGSLAFLPESFDGDPASNTAPL
QPEQLVFTLEITGTYLYISAMPDLSPLSVFQNLQVIRGRILHNGASLYLQLGLGI
SWLGRSLRELGSGLALIHNTHLCFVHTVPMWQLFRNPHQALLHTANRDEDCVGGG
LACHQICARGHCGFGPTQCVNCSQFLRGOECVEECRVLQGLPREYVNARHCLPCHPE
COPQSGSVTCFGEADQCVACAHYKDPFCVACPSGVKPDLSYMP IWKFPDEGACQ
PCP INCHSCVDLDKQCPAEORASPLTSII SAVVIGLLVVLGVFGVGLIKRROQKI
RKYTRRLLOETELVEPLTPSGAMPNQAMRILKTELKVKVLGSGAGTGYVKGWI
PQENYKIVPAIKLURENTSPRANKILDOEAIYVAGVSGPYVSRLLIGICTLSTVQLVT
QLMYGCILLDHRENRGRSLGSDLLNMCQIAKMSYLEDVLRVHRLDLAARNVLKSP
NHWKITDFGLARLLDIDETEYHADGKVP I KWMALESTILRRFTHQSDVSGVTYWE
LMTFGAKPYDGIKPAEIPDLLEKGRLPQPICTIDVYIMVKNKMSDECRPFREL
VSEFSRMDPQRFVYVIONEDLGPASPLSTFVRSLEDDMDGLVDAEVLVPOQGF
FCPDAPAGAGGVHHRHSRSTSGGDLTLGLEPSEEAAPSPAPASGAGSDVFDG
DUGMAAGQLSLPDRPQRYSEDPTPLPSETDGYVAPLTCSFPQPEYVNPQDVR
POPPSPREGPLPAARPAAGATLERPKT LSPGKNGVKVDFAFGGAVENPNLYLPQGGAA
POPHPPAPSPAFDNLVYWDQDPPERGAPEPTFKGTPTAENPEYLGLDVPV"
BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN
Alignment Scores:
Pred. No.: 5,39e-187 Length: 3768
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
Db: 6 Gaps: 1
SE06 (1-919) x AX384604 (1-3768)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
|||||
Db 1 ATGGAGCTGGCGCCTTGTGCCCTGGGGCTCTCTCTCGCCCTCTGCCCCCGGAGCC 60

QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProLaserProGlu 40
|||||
Db 61 GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCTGCGAGTCCCGAG 120

QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
|||||
Db 121 ACCACCTGGAGATGCTCCGCCACCTCTACAGGGGCTGCCAGTGTGGAGGAAACCTG 180

QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnIuVal 80
|||||
Db 181 GAACCTACCTACCTGCCACCAATGCCAGCTGTCTCTCTCGAGATATCCAGGAGGTG 240

QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
|||||
Db 241 CAGGGCTAGCTGCTCATCGCTCACACCAAGTGAGGAGGTCCCTGCTGAGAGGCTGCGG 300

QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
|||||
Db 301 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCCCTGCGCTGTAGACAATGA 360

QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
|||||
Db 361 GACCCGCTGAACAATACCAACCCCTGTGCAGGGGCTCCCGAGGAGGCTGCGGGAGCTG 420

QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
|||||
Db 421 CAGCTTCGAAGCCTCACAGAGATCTTTGAAAGAGGGGTCTTGTATCCAGGAGACCCCGAG 480

QY 161 LeuCysTyrGlnAspThrIleLeuTyrPlysAspIlePheHisLysAsnAsnGlnLeuAla 180
|||||
Db 481 CTCTGCTACCGAGACACGATTTTGTGAAGAGGGGTCTTGTATCCAGGAGACCCCGAGTGGCT 540

QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
|||||
Db 541 CTCACACTATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGCTAAG 600

QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
|||||

```





Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	301	ATTGTGGAGGACCCAGCTCTTTGAGGACAACTATCCCTGGCGGTGTAGACAATGGA	360
Qy	121	AspProLeuAsnAsnThrThrProValThrClyAlaSerProGlyGlyLeuArgGluLeu	140
Db	361	GACCCGCTGAACAATACACCCCTGTCCAGGGGCTCCCGAGAGGCTTCGGGAGCTG	420
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160
Db	421	CAGCTTGAAGCCTCACAGAGATCTTGAAGGAGGGTCTTGATCCAGCGGAACCCCGAG	480
Qy	161	LeuCysTyrGlnAspThrIleLeuTrpPlyAspIlePheHisLysAsnAsnGlnLeuAla	180
Db	481	CTCTGCTTACCAGGACACGATTTGTGGAAGGACATCTCCACAAGAACAAACACAGCTGGCT	540
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	541	CTCACACTGATAGACACCAACCGCTTCGGGGCTGCCACCCGTCTTCCGATGTGTAAAG	600
Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	601	GGTCCCGCTGTGGGAGAGATCTCTGAGGATGTCTGAGAGCTTGACGCCGACTGTCTGT	660
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	661	GCCGCTGGCTGTGCCGCTCAAGGGGCCACTGCCACTGACTGCTGCCATGAGCAGTGT	720
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	721	GCTGCCGGCTGTCAGGGGCCCAAGCACTGTGACTGCTGGCTGCCCTTCAACCCAC	780
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	781	AGTGGCATCTGTGAGCTGGACTGCCAGCCCTGGTCACTTCAACACAGACAGTTTGAG	840
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	841	TCCATGCCCAATCCGAGGGCCGATATACATTTGGGCCAGCTGTGTGACTGTGCTGCC	900
Qy	301	TyrAsnTrpLeuSerThrAspValGlySerCysThrTrpLeuValCysProLeuHisAsnGln	320
Db	901	TACAACCTACCTTCTACGGACGTGGGATCTGCACCCCTGCTGTGCCCTTGCACACCAA	960
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1020
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGCTTGGGCATGGAGCACTTCCGAGAGGTGAGGGCAGTTACACAGTGCCAAT	1080
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGAGATTTGCTGGCTCCAGAAGATCTTTGGAGCTTGGCATTTCTGCCCGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1141	TTTGATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGCAGCAGCTTCCAAAGTGTTT	1200
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGGACAGCTGCGCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTACGCTGTCTCCAGAACCCTGCAAGTAAATCCGGGGAGCAATCTGCACAATGGCGCC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCTTGCAGGGCTGGGCGATCAGCTGGCTGGGGCTGGCGCTCCTCCTGAGGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480

1381	Db		CTGGCAGTGGACTGGCCCTCATCCACATTAACACCCACCTCTGCTTCTGTCGACACGGTG	1440
481	Qy		ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
1441	Db		CCCTGGGACCAAGCTCTTTCCGGAACCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1500
501	Qy		GluAspGlnCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
1501	Db		GAGSACAGTGTGTGGCGAGGGCTTGGCCTGCCACCAAGCTGTGCGCCCGAGGCACTGC	1560
521	Qy		TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
1561	Db		TGGGGTCCAGGGGCCACCCAGTGTGTAAGTGCAGCCAGTCTCTTCTGGGGCCAGGAGTGC	1620
541	Qy		ValGluGlnCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
1621	Db		GTGAGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGCACTGT	1680
561	Qy		LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
1681	Db		TTGCGGTGCCACCTCAGTGTCAAGCCCAAGATGGTTCAGTGAAGCTGTTTGGACGGGAG	1740
581	Qy		AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
1741	Db		GCAGCACAGTGTGGGCTGTGCCCATATAAGAGACCTCCCTTCTGCGTGGGCCCGCTGC	1800
601	Qy		ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
1801	Db		CCCAGCGGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
621	Qy		GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
1861	Db		GGCGCATGCCAGGCTTGGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	1920
641	Qy		GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
1921	Db		GGTGCCCCCGCAGAGAGAGCGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGC	1980
653	Qy		-----	653
1981	Db		ATTCTGCTGCTGTGTCTTTGGGGTGGTCTTTTGGGATCTCTCATCAAGCAGCGCAGCAG	2040
653	Qy		-----	653
2041	Db		AAGATCGGAAGTACAGATGCGGAGACTCTCGAGGAACGAGCTGTGTGGAGCCGCTG	2100
653	Qy		-----	653
2101	Db		ACACCTAGCGGAGCGATGCCCAACCAAGCCACAGATGCGGATCTCTGAAAGAGCGGAGCTG	2160
653	Qy		-----	653
2161	Db		AGGAAGGTGAAGTGTGTGATCTGGCGCTTTTGGCACAGTCTACAAGGCACTCTGGATC	2220
653	Qy		-----	653
2221	Db		CCTGATGGGAGAAATGTGAAATTCAGTGGCCATCAAAAGTGTGAGGGGAAAACATCC	2280
653	Qy		-----	653
2281	Db		CCCAAGCCCAAGAAATCTTAGACGAGCATACGTGTATGGCTGTGTGGCTCCCCA	2340
653	Qy		-----	653
2341	Db		TATGTCTCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGTGTGACACAGCTT	2400
653	Qy		-----	653
2401	Db		ATGCCCTATGGCTGCCCTTTAGACCATGTCCGGGAAAACCGCGGACGCCCTGGCTCCCAAG	2460
653	Qy		-----	653







Db	595	CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGCAACCCCCAG	654	Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Qy	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180	Db	1735	TGGGGTCCAGGGGCCACCCAGTGTCAACTGCAGCCAGTTCTTCGGGGGCCAGGAGTGC	1794
Db	655	CTCTGCTACCAAGGACACCATTTGTGGGAAGGACATCTTCCCAAGAACCAACAGCTGGCT	714	Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200	Db	1795	GTGGAGGATGCCAGTACTGCAGGGGCTCCCAAGGAGTATGTGAATGCCAGGCACCTGT	1854
Db	715	CTCACACTGATAGACACCAACCGCTCTCGGGCCCTGCCACCCCTGTTCTCCGATGTGAAG	774	Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Qy	201	GlySerArgCysTrpGlyLysSerGluAspCysGlnSerLeuThrArgThrValCys	220	Db	1855	TTGCCGTGCCACCCCTGAGTGTGACGCCAGAAATGGCTCAGTGACCTGTTTGGACCGAG	1914
Db	775	GGCTCCCGCTCTCGGGAGAGAGTCTGAGGATTTGTCAGAGCCCTGACGCGCACTGTCTGT	834	Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240	Db	1915	GCTGACCAGTGTGGGCTGTGCCCACTATTAAGACCCCTCCCTTCTGCGTGGCCGCTGC	1974
Db	835	GCCGGTGCTGTGCCGTGCAAGGGCCACTGCCACTGACTGCTGCCATGACAGAGTGT	894	Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260	Db	1975	CCCAGCGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG	2034
Db	895	GCTGCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGGCTGCCTCCACTTCAACAC	954	Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280	Db	2035	GGCGCATGCCAGCCTTGCCCATCAACTGCACCCACTCTGTGTGGACCTGGATGACAAG	2094
Db	955	AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGTGTCACTAACACACAGACACTTTGAG	1014	Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300	Db	2095	GGCTGCCCGCGGACGACAGAGCCGCCCTCTGACGTCCATCATCTCTGCGGGTGGTGGC	2154
Db	1015	TCCATGCCCATTCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGTGCC	1074	Qy	653	-----	653
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320	Db	2155	ATTCTGCTGCTGCTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGACGCGCAG	2214
Db	1075	TACAACCTACCTTTCTACGAGCGTGGATCTCTGCACCCCTCGTCTGCCCTTSCACACCA	1134	Qy	653	-----	653
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340	Db	2215	AAGATCCGGAAGTACACGATCCGGAGACTGCTGCGAGAAACGGAGCTGGTGGAGCCGCTG	2274
Db	1135	GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1194	Qy	653	-----	653
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	Db	2275	ACACCTAGCGGAGCGATGCCCAACAGCGGAGATCGGGATCTCTGAAGAGACGAGGCTG	2334
Db	1195	GTGTGCTATGCTTGGCATGAGCAGCTTGGCAGAGGTGAGGCGATTTACAGTGGCCAAT	1254	Qy	653	-----	653
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380	Db	2335	AGGAAGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2394
Db	1255	ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGAGCCCTGGCATTTCTGCCGAGAGC	1314	Qy	653	-----	653
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400	Db	2395	CCTGATGGGAGAAATGTGAATAATTCAGTGGCCATCAAAAGTGTGTGAGGGAACACATCC	2454
Db	1315	TTTGATGGGACCCAGCCTCCAAACACTGCCCGCTCCAGCCAGACAGCAGCTCCAAGTGT	1374	Qy	653	-----	653
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro	420	Db	2455	CCCAACCCCAACAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGGGCTCCCCA	2514
Db	1375	GAGACTCTGGGAAGATCACAGTTACTATACATCTCAGCATTCGCGGAGCGGCTGCCT	1434	Qy	653	-----	653
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440	Db	2515	TATGCTCTCCCGCTTCTTGGCATCTGCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2574
Db	1435	GACCTCAGCGCTTCCAGAACCTCAAGTAATCCGGGACGAAATCTGCACAAATGGCGCC	1494	Qy	653	-----	653
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460	Db	2575	ATGCCCTATGGCTGCCTTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCAG	2634
Db	1495	TACTCGCTGACCTTCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGAA	1554	Qy	653	-----	653
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480	Db	2635	GACCTGCTGAACCTGGTGTATGACAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG	2694
Db	1555	CTGGCAGTGACTGGCCCTCATCCACATTAACACCACCTCTCTCTCGTGACACAGGTG	1614	Qy	653	-----	653
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500	Db	2695	CTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGTGCTCAAGAGTCCCAACCATGTCAA	2754
Db	1615	CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGTCTCCACACTGCCAACCGGCCA	1674	Qy	653	-----	653
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520	Db	2755	ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTTCACGAGACAGAGTACCATGCAGAT	2814
Db	1675	GAGCAGAGTGTGTGGGCGAGGGCTGGCCCTGCCACCAAGCTGTGCGGCCCGGAGGCACTGC	1734				

```
QY 653 ----- 653
Db 2815 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCGCGCGGTTTCACC 2874
QY 653 ----- 653
Db 2875 CACCAGAGTGTGTGTGGAGTTATGTGTGCTGCTGAGCTGATGACTTTTGGGGCC 2934
QY 653 ----- 653
Db 2935 AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAAAGGGGAGCGG 2994
QY 653 ----- 653
Db 2995 CTGCCCCAGCCCCCATCTGCACCAATTGATGTCTACATCATGTGCTCAAAATGTTGGATG 3054
QY 653 ----- 653
Db 3055 ATTGACTCTGAATGTCGGCCAAAGATTCCGGGAGTTGGTGTCTGAATTCCTCCGCAATGGCC 3114
QY 654 ----- 654
Db 3115 AGGGACCCCGCGCTTTGTGTCATCCAGAAATGAGACTTGGGCCCCACCCAGTCCCTTG 3174
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3175 GACAGCACCTTCTACCGCTCACTGTGTGGAGGACGATGACATGGGGACCTGGTGGATGCT 3234
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3235 GAGGAGTATCTGGTACCCAGCAGGCGCTTCTTCTGTCCAGACCTGCCCGCGCGCTGGG 3294
QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724
Db 3295 GGCATGCTCCACCACAGGACCGCAGCTCATCTACAGAGAGTGGCGGTGGGACCTGAC 3354
QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744
Db 3355 CTAGGCTGGAGCCCTCTGAAGAGGAGGCCCGCCAGCTCTCCACTGSCACCTCCGAAGG 3414
QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeuGlnSer 764
Db 3415 GTGGCTCGATGTATTTGATGGTACCTGGGAATGGGGGAGCCCAAGGGCTGCAAGC 3474
QY 765 LeuProThrHisAspProSerProLeuGlnAlaProTyrSerGluAspProThrValProLeu 784
Db 3475 CTCCCCACATGACCCCGACCTCTACAGCGGTACAGTACAGACCCACATGACCTCC 3534
QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804
Db 3535 CCTCTGAGACTGATGGCTACGTTGCCCCCTGACCTGCAGCCCCCAGCCTGAATATGTG 3594
QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824
Db 3595 AACACCCAGATGTTGGCCCCCAGCCCCCTTCGCCCCGAGAGGGCCCTCTGCTGCTGCC 3654
QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844
Db 3655 CGACCTGCTGGTGGCCACTTGAAAGGCCACAGACTCTCTCCCGAGGAGAAATGGGGTC 3714
QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864
Db 3715 GTCAAGAGAGTTTTCCTTTGGGGTGGCGTGGAGAACCCCGAGTACACCCCCAG 3774
QY 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884
Db 3775 GGAGGAGCTGCCCCCTCAGCCCCCAGCCCCCTCTCTGCTGCTTCCAGCCCGCTTCACAACTC 3834
QY 885 TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904
Db 3835 TATTACTGGACGACGACCCACAGAGCGGGGGCTCCACCCAGACCTTCAAGAGGACA 3894
QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
```

```
Db 3895 CCTAGGCGAGAAACCCAGAGTACCTGGCTGTGGAGTCCAGTG 3939
RESULT 7
LOCUS AR167390
DEFINITION Sequence 26 from patent US 6287569.
ACCESSION AR167390
VERSION AR167390.1 GI:17903168
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4473)
AUTHORS Kips, T.J. and Wu, Y.
TITLE Vaccines with enhanced intracellular processing
JOURNAL Patent: US 6287569-A 26 11-SEP-2001;
FEATURES
SOURCE Location/Qualifiers
1..4473
/organism="unknown"
BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN
Alignment Scores:
Pred. No.: 6,37e-187 Length: 4473
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 6 Gaps: 1
SEQ6 (1-919) x AR167390 (1-4473)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 175 ATGGAGCTGGCGGCTTGTGGCGCTGGGGCTCTCTCTCGCCCTCTTGTCCCGGAGCC 234
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCGAGCACCAAGTGTGCACCGCACACATGAAGTGGCTCGCTCCAGTCCCGAG 294
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 295 ACCACCTGGACATGCTCGCCACCTCTACAGGGCTGCAGGTGTGCAGGAAACCTG 354
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 355 GAACCTACCTACCTGCCCAATGCCAGCTGTCTTCTCTGCAGGATATCCAGGAGTG 414
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGGCTAGTGTCTATCGCTCACACCAAGTGCAGGAGTCCCACTGCAGAGGTGCGG 474
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 475 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGTGCGCTGTACACAATGA 534
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGCTGAACAATACCACTCCCTGTGCAGGGGCTTCCCGAGGAGCTCGGGGAGCTG 594
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 595 CAGCTTCGAGGCTCACAGAGATCTTGAAGGAGGGGTCTTGTATCCAGCGGAACCCAG 654
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
Db 655 CTCTGCTACCGAGACACGATTTTGTGGAAGGACATCTTCCACAAGAAACAACAGCTGGCT 714
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 715 CTCACACTGATAGACACCAACCGCTCTCGGGCTTGCACCCCTGTCTTCCGATGTGTAAG 774
```

Qy	201	GlySerArgCysTrpGlyGlyLeuSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	775		834
		GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTGAGAGCCTGACGCGCACTGTGTGT	
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	835		894
		GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT	
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	895		954
		GCTGCCGGCTGCACGGSCCCCAAGCACTCTGACTGCTGCCCTGCCCTCCACTTCAACCAC	
Qy	261	SerGlyIleCysGluLeuLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	955		1014
		AGTGGCATCTGTGAGCTGTCACCTGCCAGCCCTGGTCACTTACAAACAGACACAGTTTGAG	
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	1015		1074
		TCCATGCCCAATCCCGAGGGCCGGTATACATTTGGCGCCAGCTGTGAGCTGCTGTGCC	
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	1075		1134
		TACAACATACCTTTCTACGGAGCTGGGATCTGCACCCCTGCTGCCCCCTGCACAACCAA	
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	1135		1194
		GAGGTGACAGCAGAGGATGGAACACACACGGGTGTGAGAAGTGCAGAAAGCCCTGTGCCCGA	
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1195		1254
		GTGTGCTATGGTCTGGGCATGGAGCACTTCGAGAGGTGAGGGCAGTTACAGTGCCCAAT	
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1255		1314
		ATCCAGAGATTGCTGGCTGCAGCAAGATCTTTGGGAGCCTGGCATTTCTCCCGGAGAGC	
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1315		1374
		TTTGATGGGACCCAGCCTCCAACTATGCCCGCTCCAGCCAGAGCAGCTCCAAAGTGT	
Qy	401	GluThrLeuGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1375		1434
		GAGACTTGAAGAGATCACAGGTACCTATACATCTCAGCATGGCCGAGACGCCCTGCC	
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1435		1494
		GACCTCAGCGTCTCCAGAACCTGCAAGTAATCCGGGAGCAATTTGTCACATGGCGCC	
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1495		1554
		TACTCGGTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACGTAGGGAA	
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1555		1614
		CTGGGCAGTGGAGCTTCATCCACATTAACCCACCTCTGCTTCGTGTCACACGGTG	
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1615		1674
		CCCTGGGACCAAGCTCTTTCCGAGACCCGACCAAGACTCTGCTCCACACTGCCAACCGGCCA	
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1675		1734
		GAGGACAGTGTGGGCGAGGGCCTGGCCTGCCACCAAGCTGTGCGCCCGAGGGCACTGC	
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1735		1794
		TGGGGTCCAGGGGCCACCCAGTGTGTAACCTGAGCCAGTTCCTTCGGGGCCAGGAGTGC	
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1795		1854
		GTGAGGAATGCCAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580

Db	1855	 TTGCCGTGCCACCCTGAGTGTACGCCCCAGAAATGGCTCAGTGCACCTGTTTGGACCGGAG	1914
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys	600
Db	1915	 GCTGACCAGTGTGGCTGTGCCCATATATAGAGACCTCCCTTCTCGGTGGCCCGCTGC	1974
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1975	 CCCAGCGGTGTGAACCTGACCTCTCCTACATGCCCATCTGGAAAGTTTCCAGATGAGGAG	2034
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2035	 GGCGCATGCCAGCCTTGCCCATCACTGCACCCACCTCCTGTGTGGACCTGGATGACAAG	2094
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2095	 GGTGCTCCCGCGGACGACGAGAGCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTTGGC	2154
Qy	653	----- -----	653
Db	2155	ATTCGTCTGTCGTGGTCTTTGGGGTGGTCTTTGGGATCCTCATCAAGACGACGGCAGCAG	2214
Qy	653	----- -----	653
Db	2215	AAGATCCGGAAGTACACGATCCGGAGACTGCTGCGAGGAACGGAGCTGTTGGAGCCCGTG	2274
Qy	653	----- -----	653
Db	2275	ACACTAGCGGAGCGATGCCCAACGCGCGAGATGCGGTAGGAAACGGAGCTGTTGAGCGCGCTG	2334
Qy	653	----- -----	653
Db	2335	AGGAAGGTGAAGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2394
Qy	653	----- -----	653
Db	2395	CCTGATGGGAGAAATGTGAATAATTCACAGTGGCCATCAAGTGTGAGGGGAAAACACATCC	2454
Qy	653	----- -----	653
Db	2455	CCCAAGCCACAAAGAAATCTTAGACGAAGCATACGTGTGGTGGTGGGTCTCCCCA	2514
Qy	653	----- -----	653
Db	2515	TATGTCTCCCGCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2574
Qy	653	----- -----	653
Db	2575	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCCGGACGCCCTGGGTCTCCAG	2634
Qy	653	----- -----	653
Db	2635	GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGCGG	2694
Qy	653	----- -----	653
Db	2695	CTCGTACACAGGGACTTTGGCGCTCGGAACGTGCTGTCAAGAGTCCCAACCATGTCAAA	2754
Qy	653	----- -----	653
Db	2755	ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTCACGACAGAGTACCATGCAGAT	2814
Qy	653	----- -----	653
Db	2815	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCCGCGCGGTTCACC	2874
Qy	653	----- -----	653
Db	2875	CACCAGAGTGATGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2934
Qy	653	----- -----	653

Db 2935 AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2994  
 QY 653 ----- 653  
 Db 2995 CTGCCCCAGCCCCCATCTGCACCATTTGATGCTACATGATCATGTCAATGTTGGATG 3054  
 QY 653 ----- 653  
 Db 3055 ATTGACTCTGAATGTCCGCCAAGATTCGCGGAGTTGGTGTCTGAATTCCTCCCGCATGGCC 3114  
 QY 654 ----- 654  
 Db 3115 AGGACCCCGACCGCTTTGTGTCTATCCAGATGAGGACTTGGCGCCAGCCAGTCCCTTG 3174  
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684  
 Db 3175 GACAGCACTTCTACCGCTCACTGTGTGAGGAGGATGACATGGGGACCTGGTGGATGCT 3234  
 QY 685 GluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGlyAlaGly 704  
 Db 3235 GAGGAGTATCTGTATCCAGCAGGCGCTTCTCTGTCCAGACCTGCCCGCGCGCTGGG 3294  
 QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724  
 Db 3295 GGCATGCTCCACACAGGACCGACCTCATCTACAGGAGTGGCGGTGGGACCTGACA 3354  
 QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744  
 Db 3355 CTAGGCTGGAGCCCTCTGAGAGGAGGCCCCAGGTCTCCACTGCGACCTCCGAGGG 3414  
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764  
 Db 3415 GCTGGCTCGATGATTTGATGTGACCTGGGAATGGGGAGCGCCAGGGGCTGCAAGC 3474  
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
 Db 3475 CTCCCCACATGACCCCGACCTCTACAGCGGTACAGTGAGACCCACAGTACCCCTG 3534  
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
 Db 3535 CCCTCTGAGACTGATGGCTACGTTGCCCTGACCTGCAGCCCCCAGCCCTGAATATG 3594  
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAla 824  
 Db 3595 AACACCCAGATGTTGCGCCCGACGCCCTTGCCCGCAGAGGCGCTCTGCTGCTGCTG 3654  
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
 Db 3655 CGACCTGCTGGTGCCACTCTGGAAGGCCACAGACTCTCTCCCGAGGAAGTGGGGTC 3714  
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
 Db 3715 GTCAAGACGTTTTTGCCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCAG 3774  
 QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884  
 Db 3775 GGAGGAGCTGCCCTCAGCCCGACCCCTCTCTCTGCTTCCAGCCAGCCCTTCGACACCTC 3834  
 QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
 Db 3835 TATTACTGGGACGAGACCCAGAGTACCTGGTGTGGTCTGGAGTCCAGCTTCAAGAGGACA 3894  
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
 Db 3895 CCTACGCGACAGAACCCAGAGTACCTGGTGTGGTCTGGAGTCCAGT 3939  
 RESULT 8  
 HSERB2R  
 LOCUS HSERB2R 4473 bp mRNA linear PRI 30-MAR-1995  
 DEFINITION Human c-erb-B-2 mRNA.  
 ACCESSION X03363  
 VERSION X03363.1 GI:31197  
 KEYWORDS cell surface glycoprotein; cellular oncogene; erb-B-2 cellular;

glycoprotein; growth factor receptor; kinase; neu cellular oncogene; transmembrane protein; tyrosine kinase.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 4473)  
 AUTHORS Yamamoto,T., Ikawa,S., Akiyama,T., Semba,K., Nomura,N., Miyajima,N., Saito,T. and Toyoshima,K.  
 TITLE Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth factor receptor  
 JOURNAL Nature 319 (6050), 230-234 (1986)  
 MEDLINE 86118663  
 PUBMED 3003577  
 REFERENCE 2  
 AUTHORS Papawalis,J., Nikitin,A.Yu. and Rajewsky,M.F.  
 TITLE G to A polymorphism at amino acid codon 655 of the human erbB-2/HER2 gene  
 JOURNAL Nucleic Acids Res. 19 (19), 5452 (1991)  
 MEDLINE 92020265  
 PUBMED 1681519  
 COMMENT The c-erb-B-2 protein shows similarity to the epidermal growth factor receptor.  
 FEATURES  
 source Location/Qualifiers  
 1..4473  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /strain="cell line MKN-7"  
 /db\_xref="taxon:9606"  
 175..3942  
 /note="c-erb-B-2 precursor"  
 /codon\_start=1  
 /protein\_id="CAA27060.1"  
 /db\_xref="GI:31198"  
 /db\_xref="SWISS-PROT:P04626"  
 /translation="MELALCRWGLLLALLPPGAASQTQCTGDMKRLPASPETHLD MLRLHYQGVQVQGNLELTLYPTNASLFLQIDQVQGYLHAHQVQVPLQRLIRV RTGLFEDNYALAVLDNGPLNNTPTVTGASPGRLQLRLSLTILKGLQIQRNP LCYQTLFWKIDIFHKNQNLALTLDITNRSRACPCSPMCKGSRGSESDQSLTRP VCAGCARCKGPLPTDCHECAAGCTGPKHSDCLACLFHNSGICELHCPALVTNT DTFESMPNPEGRYTFGASCVTACPYNLSTDVSGCTLYCPHUNQVETADGTCRCEK SKPCARVCYGLGMEHLREVRATISANIQFAGCKTFGLAFPLPFDGDPASNTAPL QPEQLQVETLEEITGYLYISAMPDLSPLVSFQNLQVIRGRILHNGAYSLSLQGLGI SWGLRSLRELGLALIHNTHLCTVHPWDQLFRNPQHLALTANRHPEDCEGEG LACHOLCARHCWGPGTCQVNCQFRLGQECVECRVQGLPREVYNARHCLPCHPE COPQNSVTCFGEADQCVACAHYKDPFCVACRCPGVKQDLSYMPWKFPDEGACQ PCPINCTHSCVDLDDKGPACQASPLTSLISAVVGLLVVGLVGVGLILKRRQKI RYITMRRLUQETELVEPLTPSGAMPNQAMRIILKTELKRVKVLGSGAGTGVKGIWI PDGENVKIPVAIKVIRENTSPKANKEIIDEAYVMAGVSGSPVSRLLGICLTSTVOLVT QLMPYCCLLDHYRENRLGSDLLNMCQIAKMSYLEDVRLVHRDLAARNVLKSP NHVKTDFGLARLLDIDETEYHADGKVPWKMALESILRRRTHQSDVMSGVTVWE LMTFGAKPYDGIIPAREIPDLLEKGRLPPOPICTIDVYIMVKWCMIDSECRPREL VSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDDDDMDGLVDAEYLYPQGF FCPDPAPAGGWHHRSSRSGGDLTLGLEPSEAPRSLAPSGAGSDVFDG DIGMGAAGLQSLTHDPSLPQYSEDPTPLPSTEDTVAPLTCSPQEVYNQPDVR POPPSRPGPLPAAPAGATLERPKTSLSPKNGVVKDVFAGVGAVENPEYLYPQGGAA POPHPAPAFDNLYYWDQPPERGAAPPSTFKGTPTAENPEYLLDLPV"  
 175..237  
 /note="put. signal peptide (aa -21 to -1)"  
 238..3939  
 /product="put. c-erb-B-2 protein (aa 1-1234)"  
 376..384  
 /note="pot. glycosylation site"  
 544..558  
 /note="pot. glycosylation site"  
 733..741  
 /note="pot. glycosylation site"  
 949..957  
 /note="pot. glycosylation site"  
 1762..1770  
 /note="pot. glycosylation site"  
 1885..1893  
 /note="pot. glycosylation site"  
 sig\_peptide  
 mat\_peptide  
 misc\_feature  
 misc\_feature  
 misc\_feature  
 misc\_feature  
 misc\_feature  
 misc\_feature

```

misc_feature 2059..2067
/note="pot. glycosylation site"
misc_feature 2353..3132
/note="aa 727-986, seq. homologous to EGF receptor kinase
domain"
misc_feature 2446..2454
/note="pot. glycosylation site"
misc_feature 4455..4460
/note="put. polyA signal"
polyA_site 4473
/note="polyA site"

BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN

Alignment Scores:
Pred. No.: 6.37e-187 Length: 4473
Score: 4900.00 Matches: 919
Percent Similarity: 73.23% Conservative: 0
Best Local Similarity: 73.23% Mismatches: 0
Query Match: 96.49% Indels: 336
DB: 9 Gaps: 1

SEQ6 (1-919) x HSERB2R (1-4473)

Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 175 ATGGAGCTGGCGGCTTGTGCGGCTGGGGCTCCTCTCGCCCTTGTGCCCGGAGCC 234
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCGAGCACCCAAAGTGTCCAGCGGCACAGACATGAAGCTGGGCTCCCTGCGCAGTCCCGAG 294
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 295 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGGTGGCAGGAAACCTG 354
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 355 GAATCACCCTACCTGCCACCAATGCCAGCGCTGTCCTTCCTGCGAGGATATCCAGGAGGTG 414
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGGCTACGTGCTCATCGCTCAACAAGTAGGAGGAGTCCACTGCCAGGAGCTCCGG 474
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 475 ATTGTGGAGGCCACCCAGCTCTTTGAGGACAATATGCCCTGGCCGTGCTAGACATGGA 534
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGCTGAACATACACCCCTGTTCACAGGGGCTTCCCAGAGGCTCCCGGAGCTG 594
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 595 CAGCTTCGAAGCCTCAAGAGATCTTGAAGAGGGGTCTTGATTCACAGCGGAACCCCCAG 654
Qy 161 LeuCystTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 655 CTCTGCTACCGAGACAGATTTGTGGAAGACATCTTCCACAGAACACACCGAGCTGGCT 714
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 715 CTCACATGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAG 774
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 775 GGCTCCCGCTGCTGGGAGAGATTCTGAGGATTGTGAGAGTTGTCCAGAGCCTGACCGCACTGTCTGT 834
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db 835 GCGGGTGGCTGTGCCCGCTGCAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 894
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260

```

```

Db 895 GCTGCCGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGCCCTCCACTTCAACGAC 954
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 955 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGTGTACCTTACAAACACAGACACGTTTGG 1014
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 1015 TCCATGCCCAATCCCGAGGGCCGCTATACATTCGCGCCAGCTGTGTGACTGCTGCTCC 1074
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1075 TACAACCTACCTTTCTACGGAGCTGGGATCTGCAACCTCTGCTGCTGCCCTGCACAACCA 1134
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGlnLysCysSerLysProCysAlaArg 340
Db 1135 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGGA 1194
Qy 341 ValCystTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1195 GTGTGCTATGCTGGGCATGGAGCACTTGGAGAGGTGAGGGCAGTTACCACTGCCAAT 1254
Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1255 ATCCAGCAGTTTGTGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC 1314
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
Db 1315 TTTGATGGGAGCCAGCCTCCAACACTGCCCGCTCCAGCAGAGCAGCTCAAGTGT 1374
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1375 GAGACTCTGAAGAGATCAGAGTTACTTATACATCTCAGCATGCGCGGACAGCCTGCCT 1434
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1435 GACCTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGAGCGAATCTGCACAAATGGCGCC 1494
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1495 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCTACTCAGGGAA 1554
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1555 CTGGCAGTGGACTGGCGCTCATCCACCATAAACCCACCTCTGCTTCTGTCACACCGTG 1614
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1615 CCCTGGGACCCAGCTCTTTCCGAAACCCGACCAAGCTCTGTCTCCACACTGCCAACCGGCA 1674
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1675 GAGGACAGTGTGTGGGCGAGGGCTGGCGCTGCCACAGCTGTGCGCGCCGAGGCACTGC 1734
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1735 TGGGCTCCAGGGCCACCCAGTGTGTCACTGACGCGAGTTCTTCTCGGGGCCAGGAGTGC 1794
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1795 GTGGAGGAATGCCAGTACTGTCAGGGGCTCCCGAGGGAGTATGTGAATGCCAGGCACTGT 1854
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1855 TTGCCGTGCCACCTGAGTGTGACGCCCCAGAAATGGCTCAGTGCCTGTTTGTGACCGGAG 1914
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1915 GCTGACCACTGTGGCTGTGCCACTATAGGACCTCCCTTCTGCTGCGGCGCGCTGC 1974
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620

```

Db 1975 CCACGGGTGTAACCTGACCTCTCTACATGCCCCATCTGGAAGTTTCCAGATGAGGAG 2034  
 QY 621 GlyAlaCysGlnProCysProIleasnCysThrHisSerCysValAspLeuAspAspLys 640  
 Db 2035 GGGCATGCGAGCTTGGCCCATCACTGACCCACCTCTGTGTGACCTGGATGACAAAG 2094  
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653  
 Db 2095 GGCTGCCCGCGAGCAGAGAGCCGCCCTCTGACGTCCATCATCTCTCGGTGTTGGC 2154  
 QY 653 ----- 653  
 Db 2155 ATTCTGCTGGTGGTCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGCGGAGCAG 2214  
 QY 653 ----- 653  
 Db 2215 AAGATCCGGAAGTACAGATGCGGAGACTGCTGAGGAAACGAGAGCTGGTGGAGCGGCTG 2274  
 QY 653 ----- 653  
 Db 2275 ACACCTAGCGGAGCGATGCCAACCAAGCGGAGATGCGGATCCTGAAAGAGACGAGCTG 2334  
 QY 653 ----- 653  
 Db 2335 AGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2394  
 QY 653 ----- 653  
 Db 2395 CCTGATGGGGAGATGTGAAATTCAGTGGCCATCAAGTGTGAGGAAACACATCC 2454  
 QY 653 ----- 653  
 Db 2455 CCCAAAGCAACAAGAAATCTTAGACGAAGCATACGTGATGCTGGTGTGGCTGCCCA 2514  
 QY 653 ----- 653  
 Db 2515 TATGTCTCCCGCTTCTGGGCATCTGCCTGACATCCACGGTGACAGCTT 2574  
 QY 653 ----- 653  
 Db 2575 ATGCCCTATGGTCTCTTAGACCATGTCCGGAAACCGGAGCGCTGGGTGCCAG 2634  
 QY 653 ----- 653  
 Db 2635 GACCTGCTCACTGGTGTATGAGATGTGCAAGGGATGAGTACTCTGGAGGATGTGGG 2694  
 QY 653 ----- 653  
 Db 2695 CTCGTACAGGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAA 2754  
 QY 653 ----- 653  
 Db 2755 ATTACAGACTTCGGGTGCTCGGCTGCTGGAGATTGACGACAGATACCATGCAGAT 2814  
 QY 653 ----- 653  
 Db 2815 GGGGCAAGGTGCCATCAAGTGGATGGGCTGGAGTCCATTCTCGCGCGGCTCACCC 2874  
 QY 653 ----- 653  
 Db 2875 CACCAGAGTATGTGTGGAGTTATGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC 2934  
 QY 653 ----- 653  
 Db 2935 AAACCTTACGATGGATGCCAGCCCGGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2994  
 QY 653 ----- 653  
 Db 2995 CTGCCCAAGCCCATCTGCACCATGATGATGTATACATGATCATGTGCAATGTTGGATG 3054  
 QY 653 ----- 653  
 Db 3055 ATTGACTCTGAATGTGGCCAAAGATTCGGGAGTTGGTGTCTGAATTTCTCGCATGGCC 3114

QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664  
 Db 3115 AGGAGCCCCACCGCTTGTGTGTCATCCAGAAATGAGACTTGGGGCCACCCAGTCCCTTG 3174  
 QY 665 AspSerThrPheTyrArgSerLeuGluAspAspMetGlyAspLeuValAspAla 684  
 Db 3175 GACAGCACCTTCTACCGCTCACTGTGGAGGACGATGACATGGGACCTGTGTGATGCT 3234  
 QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704  
 Db 3235 GAGGAGTATCTGTTACCCAGCAGGCTTCTTCTGTCCAGACCTGCCCGCGCGTGGG 3294  
 QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr 724  
 Db 3295 GCATGGTCCACAGGACCGCAGCTCTACAGGAGTGGGGTGGGACCTGACA 3354  
 QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744  
 Db 3355 CTAGGCTGGAGCCCTCTGAAGAGGAGGCCCCAGGTCTCCACTGGCACCTCCGAAGG 3414  
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyGlyLeuGlnSer 764  
 Db 3415 GCTGCTCCGATGATTTGATGGTACCTGGGAATGGGGGAGCCCAAGGGGCTCAAGG 3474  
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspPrcThrValProLeu 784  
 Db 3475 CTCCCCACATGACCCCGCCCTCTACAGCGGTACAGTGGAGACCCACACAGTACCCTG 3534  
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
 Db 3535 CCCTCTGAGACTGATGGTACGTTGCCCTTACCTGCAGCCCGCAGCTGCAATATGTG 3594  
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824  
 Db 3595 AACCGACGAGATGTGGCCCGCCAGCCCTTCCCGCCGAGAGGGGCTCTGCTGCTGCC 3654  
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
 Db 3655 CGACCTGTGTGGTCCACTCTGGAAGGCCACAGACTCTCCCCAGGGAGAAATGGGCTC 3714  
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
 Db 3715 GTCAAAGAGCTTTTGGCTTTGGGGTGGCGTGGAGAACCCGAGTACTGTACACCCAG 3774  
 QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884  
 Db 3775 GGAGGAGTGGCCCTCAGCCCGCCCTCTCTGCTTCCAGCCCGCTTCGACACCTC 3834  
 QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
 Db 3835 TATTACTGGGACGAGGCCACAGAGGGGGGCTCCACCCAGCAGCCTTCACACGGGACA 3894  
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
 Db 3895 CCTACGGCAGACAGCCAGAGTACCTGGGTCTGGAGCTGCCAGTG 3939  
 RESULT 9  
 AX060704  
 LOCUS  
 DEFINITION Sequence 2 from Patent WO0100244.  
 ACCESSION AX060704  
 VERSION AX060704.1 GI:12406101  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 ERICKSON, S. and Schwall, R.  
 Methods of treatment using anti-erbB antibody-maytansinoid  
 conjugates  
 Patent: WO 0100244-A 2 04-JAN-2001;  
 JOURNAL



QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	-----	653	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684	
Db	1921	GGCTGCCCGCCGAGCAGAGCCAGCCCTCTGACGTCCATCGTCTCTCGGGTGTGGC	1980	-----	3001	GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGGGTGGATGCT	3060
QY	653	-----	-----	653	-----	-----	704
Db	1981	ATTCTGCTGGTGGTCTTTGGGGTGGTCTTTTGGGATCCTCATCAAGCGAGCGCAGCAG	2040	-----	3061	GAGGAGTACTGTGTACCCAGCAGGGCTTCTTCTGTCCAGACCTGCCCGGGCGCTGGG	3120
QY	653	-----	-----	653	-----	-----	724
Db	2041	AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGAGCTGGTGGAGCCGCTG	2100	-----	3121	GSCATGGTCCACACAGCAGCCAGCTCATCTACCAGGAGTGGCGGTGGGACCTGCACA	3180
QY	653	-----	-----	653	-----	-----	744
Db	2101	ACACCTAGCGGAGCGATGCCAACACGAGCGCAGATGCGGATCCTGAAAGAGAGCGAGCTG	2160	-----	3181	CTAGGCGTGGAGCCCTCTGAAGAGGAGGCCCCAGGTCTCCACTGGCACCTCCGAAGGG	3240
QY	653	-----	-----	653	-----	-----	764
Db	2161	AGGAAGGTGAAGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGCGATCTGGATC	2220	-----	3241	GCTGGCTCCGATGTATTGTATGGTGACCTGGGAATGGGGGAGCCCAAGGGGTGCAAGC	3300
QY	653	-----	-----	653	-----	-----	784
Db	2221	CCTGATGGGGAATGTGAAATTCACGTGGCCATCAAAGTGTGAGGGAACACATCC	2280	-----	3301	CTCCCCACACATGACCCAGCCCTCTACAGCGGTACAGTGAGGAGCCCCACAGTACCCCTG	3360
QY	653	-----	-----	653	-----	-----	804
Db	2281	CCCAAGCCCAACAAGAAATCTTAGACGAAGCATACGTGATGGTGGTGGCTCCCA	2340	-----	3361	CCCTCTGAGACTGATGGTACGTTGCCCTTGCCTGACCTGCAGCCCCCAGCTGAATATGTG	3420
QY	653	-----	-----	653	-----	-----	824
Db	2341	TATGTCTCCCGCTTCTGGGCATCTGCCTGACATCCACGGTGACGTGGTGACACAGTT	2400	-----	3421	AACCAAGCAGATGTTGGCCCCAGCCCTTTCGCCCGAGAGGGCCCTCTGCCTGTGTGCC	3480
QY	653	-----	-----	653	-----	-----	844
Db	2401	ATGCCCTATGGTGCCTCTTAGACCATGTCCGGGAAACCGCGACGCTGGGTGCCAG	2460	-----	3481	CGACCTGCTGGTGCCACTCTCGAAAGGCCCAAGACTCTCTCCCAGGGAAGATGGGGTC	3540
QY	653	-----	-----	653	-----	-----	864
Db	2461	GACCTGCTGAAGTGTATGCAGATTGCCAAGGGATGACGTACCTGGAGGATGTGCGG	2520	-----	3541	GTCAAAGACGTTTTTGGCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGCACACCCAG	3600
QY	653	-----	-----	653	-----	-----	884
Db	2521	CTCGTACACAGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAA	2580	-----	3601	GGAGGAGTGGCCCTCAGCCCCACCTCTCTCTGCTTCAGCCCAAGCTTCGACAACTC	3660
QY	653	-----	-----	653	-----	-----	904
Db	2581	ATTACAGACTTCGGGCTGGCTGGCTGGACATTGACGACAGAGTACCATGCAGAT	2640	-----	3661	TATTACTGGGACGAGCCACAGAGCGGGGGCTCCACCCAGCACCTTCAAGGGGACA	3720
QY	653	-----	-----	653	-----	-----	919
Db	2641	GGGGCAAGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGGTTCACC	2700	-----	3721	CCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG	3765
QY	653	-----	-----	653	-----	-----	
Db	2701	CACCAGAGTATGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2760	-----	RESULT 10		
QY	653	-----	-----	653	-----	AX467229	
Db	2761	AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG	2820	-----	LOCUS	3768 bp	DNA linear PAT 16-JUL-2002
QY	653	-----	-----	653	-----	DEFINITION	Sequence 3 from Patent WO0234287.
Db	2821	CTGCCCCAGCCCCATCTGCACCAATTGATGTCTACATGATCATGTGTAATGTTGGATG	2880	-----	ACCESSION	AX467229	
QY	653	-----	-----	653	-----	VERSION	AX467229.1 GI:21900511
Db	2881	ATTGACTCTGAATGTCGGCAAGATTCCGGGAGTTGGTGTCTGAATTCCTCCGCGATGGCC	2940	-----	KEYWORDS		
QY	654	-----	-----	654	-----	SOURCE	
Db	2941	AGGAGCCCCAGCGCTTGTGGTCTACCAAGATGAGGACTTGGGCCCGCCAGCGTCCCTTG	3000	-----	ORGANISM		
						Homo sapiens (human)	
						Homo sapiens	
						Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
						Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.	
					REFERENCE		
					AUTHORS	Beier, A.M., Gautam, A. and Mouritsen, S.R.	
					TITLE	Novel therapeutic vaccine formulations	
					JOURNAL	Patent: WO 0234287-A 3 02-MAY-2002;	
						Pharmexa A/S (DK)	
					FEATURES		
						Location/Qualifiers	
						1..3768	
						/organism="Homo sapiens"	
						/mol_type="genomic DNA"	
						/db_xref="taxon:9606"	



[illegible]

Db	1561	TGGGGTCCAGGGCCCCCAGTCAGTGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC	1620
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGATGCCAGTACTGCAAGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1680
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCCTCAGTGTCCAGCCCAAGATGGCTCAGTACCTGTTTGGACCCGAG	1740
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCACTGTGGCCCTGTGCCCACTATAGAACCCTCCCTTCTGCGTGGCCGCTGC	1800
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTCGGAAGTTTCCAGATGAGGAG	1860
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGCATGCCAGCTTGGCCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACAG	1920
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCCGAGCAGAGACGACCCCTCTGACGTCCATCTCTCTGCGGTGGTTGGC	1980
QY	653	-----	653
Db	1981	ATTCTCTGGTCTGTGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGGACGAG	2040
QY	653	-----	653
Db	2041	AAGATCCGGAAGTACAGATGCGGACACTCTGCAGAAACGAGCTGGTGAGCCGCTG	2100
QY	653	-----	653
Db	2101	ACACCTAGCGGAGCGATGCCAACAGCGGCGAGATCGGATCCTGAAAGAGACGGAGCTG	2160
QY	653	-----	653
Db	2161	AGGAAGTGAAGTGTGGNYCTGGCTGGCTTTTGGCACAGTCTACAAAGGCGATCTGGATC	2220
QY	653	-----	653
Db	2221	CCTGATGGGAGAAATCTGAAATTTCCAGTGGCCATCAAAGTGTGAGGGAACACATCC	2280
QY	653	-----	653
Db	2281	CCAAAGCCAACAAGAAATCTTTAGACGAAGCATACGTGATGGCTGGTGGGCTCCCA	2340
QY	653	-----	653
Db	2341	TATGTCTCCGCCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2400
QY	653	-----	653
Db	2401	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCAG	2460
QY	653	-----	653
Db	2461	GACCTCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGCTTACCTGGAGGATGTGCGG	2520
QY	653	-----	653
Db	2521	CTCGTACACAGGACATTGGCCGCTCGAAGCGTCTGGTCAAGAGTCCCAACCATGTCAA	2580
QY	653	-----	653
Db	2581	ATTACAGACTTCGGGCTGGCTGGCTGTGGACATTGACGAGACAGTACCATGACAGAT	2640
QY	653	-----	653
Db	2641	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTTCTCGCGCGGCTTCAAC	2700
QY	653	-----	653
Db	2701	CACCAGTCAATGTGTGGATTATGTGTCACTGTGTGGAGCTGATGACTTTTGGGGCC	2760
QY	653	-----	653
Db	2761	AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG	2820
QY	653	-----	653
Db	2821	CTGCCCCAGCCCCCATCTGCACCAATTGATGTCTACATGATCATGTCAAATGTTGGATG	2880
QY	653	-----	653
Db	2881	ATTGACTCTGAATGTCCGCCAAGATTCCCGGAGTTGGTGTCTGAATTTCTCCCCATGGCC	2940
QY	654	-----	654
Db	2941	AGGGACCCCGCAGCGCTTTGTGGTCAATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTG	3000
QY	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACCTTCTACCGCTCACTGCTGGAGGACATGACATGGGGACCTGTGTGATGCT	3060
QY	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3061	GAGGATATCTGTATCCCGCAGCAGGCTTCTCTGTCCAGACCTGCCCGGCGCTGGG	3120
QY	705	GlyMetValHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeuThr	724
Db	3121	GGCATGGTCCACCACAGCCAGCCAGCTCATCTACCAGGAGTGGCGTGGGGACCTGACA	3180
QY	725	LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly	744
Db	3181	CTAGGCTCGAGCCCTCTGAAGAGGAGGCCCTCCAGTCTCCACTGGCACCTCCGAAGG	3240
QY	745	AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer	764
Db	3241	GCTGGCTCGATGATTTGATGTGACTGGGAAATGGGGCAGCCCAAGGGGCTGCAAGC	3300
QY	765	LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu	784
Db	3301	CTCCCCACACATGACCCAGCCCTCTACAGCGGTACAGTGAAGACCCACAGTACCCCTG	3360
QY	785	ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal	804
Db	3361	CCCTCTGAGACTGATGGTACGTTGCCCTCCCTGACCTGCAGCCCCCAGCTGAATATGTG	3420
QY	805	AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla	824
Db	3421	AACCCAGCAGATGTTCCGCCCCAGCCCTTCCGCCCGAGAGGGCCCTCTGCCTGCTGCC	3480
QY	825	ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal	844
Db	3481	CGACCTGCTGGTCCCACTCTGGAAGGGCCCAAGACTCTCTCCCCAGGAGAAATGGGTC	3540
QY	845	ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln	864
Db	3541	GTCAAAGAGCTTTTTCCTTTGGGGTGGCGTGGAGAACCCCGAGTGTGACACCCAG	3600
QY	865	GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu	884
Db	3601	GGAGGAGCTGCCCTCAGCCCCACCTCTCTCTCCCTTCCAGCCAGCCCTTCGACACCTC	3660
QY	885	TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr	904
Db	3661	TATTACTGGACCCAGCACCCAGAGCGGGGCTCCACCCAGCACCTTCAAAGGACA	3720
QY	905	ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal	919
Db	3721	CTTACGGCAGAGAACCCAGAGTACCTGGGTCTGGAGTCCATTTCTCGCGCGGCTTCAAC	3765

## RESULT 11

AX481438  
 LOCUS AX481438 3768 bp DNA linear PAT 16-AUG-2002  
 DEFINITION Sequence 52 from Patent WO02055693.  
 ACCESSION AX481438  
 VERSION AX481438.1 GI:22316352  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 JOURNAL

Kreutzer R., Limmer S., Rost S. and Hadwiger P.  
 Method for inhibiting the expression of a target gene  
 Patent: WO 02055693-A 52 18-JUL-2002;  
 Ribopharma AG (DE)

## FEATURES

source Location/Qualifiers  
 1..3768  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 758 a 1170 c 1121 g 719 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 1.12e-186 Length: 3768  
 Score: 4892.00 Matches: 918  
 Percent Similarity: 73.15% Conservative: 0  
 Best Local Similarity: 73.15% Mismatches: 1  
 Query Match: 96.34% Indels: 336  
 DB: 6 Gaps: 1

SEQ6 (1-919) x AX481438 (1-3768)

Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20  
 Db 1 ATGGAGCTGGCGGCTTGTGCGGCTGGGGCTCTCTCTGCGGCTTTGCCCGGAGCC 60  
 Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 Db 61 GCGAGCACCCAGTGTCCACCGGCACAGATGAAGCTGGGCTCCCTGCGAGTCCCGAG 120  
 Qy 41 ThrHisLeuAspMetLeuArgHisLeuTrpGlnGlyCysGlnValValGlnGlyAsnLeu 60  
 Db 121 ACCACCTGGACATGCTCGGCCACTCTACAGGCTGCCAGGTGGTGCAGGGAACCTG 180  
 Qy 61 GluLeuThrTrpLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
 Db 181 GAATCACCCTACCTGCCCGACCAATGCCAGCTGTCTTCTGCGAGGATATCCAGGAGGTG 240  
 Qy 81 GlnGlyTrpValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 Db 241 CAGGGCTACGTGCTCATCGCTCAACAAGTAGGAGGCTGCCACTGCGAGGCTCGGG 300  
 Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAlaValLeuAspAsnGly 120  
 Db 301 ATTGTGGAGGACCCAGCTCTTGGAGCAACTATGCCCTGGCGTGTAGACAAATGA 360  
 Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 Db 361 GACCCGTGAACATACACCCCTGTTCACAGGGGCTTCCACAGGAGGCTTCCGCGGAGCTG 420  
 Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160  
 Db 421 CAGCTTCGAAGCTTCAGAGATCTTGAAGGAGGCTCTGTATCCAGCGGAACCCCGAG 480  
 Qy 161 LeuCysTrpGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180  
 Db 481 CTCTGCTACCGAGACAGATTTGTGGAAGACATCTTCCACAGACACACAGCTGGCT 540  
 Qy 181 LeuThrLeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTTCCAGGCTGCCACCCCTGTTCTCCGATGTGAAG 600

Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 Db 601 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGAGTTGTGAGAGCTTCAGAGCCTGAGCGGACTGTCTGT 660  
 Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
 Db 661 GCCGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCACTGACTGCTGCCATGAGCAGTGT 720  
 Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 Db 721 GCTCCGGCTGCAGGGCCCAAGCACTCTGACTGCTGGCTGCTCCCTCCACTTCAACCA 780  
 Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTrpAsnThrAspThrPheGlu 280  
 Db 781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGCTACCTACACACAGACACGTTTGA 840  
 Qy 281 SerMetProAsnProGluGlyArgTrpPheGlyAlaSerCysValThrAlaCysPro 300  
 Db 841 TCCATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGACTGCTGCTGCC 900  
 Qy 301 TyrAsnTrpLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 Db 901 TACAACCTACCTTCTACGGAGCTGGGATCCCTGCCACCCTGCTGCCCGCTCCACCA 960  
 Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 Db 961 GAGGTGACAGCAGAGGATGGAACACACGCGTGTGAGAAGTGCACAGACCCCTGTGCCGA 1020  
 Qy 341 ValCysTrpGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 Db 1021 GTGTGCTATGCTCTGGGCATGGAGCACTTGGAGAGGTGGAGGCGAGTTACCAATGCCA 1080  
 Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 Db 1081 ATCCAGGAGTTGTGTGCTGCAAGAAGATCTTTGGAGCCTTGGCATTTCTGCCGAGAGC 1140  
 Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
 Db 1141 TTTGATGGGACCCAGCCTCCACACTGCCCGCTCCAGCCAGAGAGCTCCAAGTGT 1200  
 Qy 401 GluThrLeuGluGluIleThrGlyTrpLeuTrpIleSerAlaTrpProAspSerLeuPro 420  
 Db 1201 GAGACTCTGGAAGAGATCAGAGTTTACCTATACATCTCAGCATGCGCGGACAGCCTGCCT 1260  
 Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 Db 1261 GACCTCAGCGCTCTTCCAGAACCTGCAAGTAAATCCGGGGAGCAATTTCTGCAATGGG 1320  
 Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 Db 1321 TACTCGCTGACCTTGCAGGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGAA 1380  
 Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480  
 Db 1381 CTGGCAGTGGAGCTGGGCTTCCAGAACCCGACCAAGCTGCTGCCACTGCGCAACCGGCA 1440  
 Qy 481 ProTrpaspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 Db 1441 CCCTGGACCAAGCTCTTTCGGAACCCGACCAAGCTGCTGCCACTGCGCAACCGGCA 1500  
 Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 Db 1501 GAGGACGAGTGTGGGCGAGGGCTGGGCTGCCAGCAGCTGTGCGCGCGGAGGCACTGC 1560  
 Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 Db 1561 TGGGCTCCAGGGCCCAAGCTGTGTCACTGACAGCAGTTCCTTCCGGGCGGAGGAGTGC 1620  
 Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTrpValAsnAlaArgHisCys 560  
 Db 1621 GTGGAGGAATGCCGAGTACTGTCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680

QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 Db 1681 TTGCCGTGCCACCTTGAGTGTGAGCCAGATGCTCAGTACCTGTTTGGACCGGAG 1740  
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
 Db 1741 GCTGACCACTGTGGCGCTGTGCCCACTATAAGGACCTCCCTTCTGCGTGGCGGCTGC 1800  
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
 Db 1801 CCAGCGGTGTGAACCTTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860  
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640  
 Db 1861 GCGCATGCCAGCCTTGCCCATCACTGCACCCCACTCTCTGTGTGACCTGGATGACAG 1920  
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer- 653  
 Db 1921 GGCTGCCCGCGGAGCAGAGACGACCTCTGAGGTCCATCGTCTCGGTGGTGGC 1980  
 QY 653 ----- 653  
 Db 1981 ATTCTGCTGGTGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGGAGCAG 2040  
 QY 653 ----- 653  
 Db 2041 AAGATCCGGAAGTACACGATGCGAGACTGCTGCAGAAACGGAGCTGTGTGGAGCGCTG 2100  
 QY 653 ----- 653  
 Db 2101 ACACCTAGCGGAGCGATGCCAACCGCGCAGATGCGGATCCTGAAGAGACGAGCTG 2160  
 QY 653 ----- 653  
 Db 2161 AGAAGGTGAAGTGTGGATCTGGCGTTTGGCACAGTCTACAAAGGGCATCTGGATC 2220  
 QY 653 ----- 653  
 Db 2221 CCTGATGGGAGAAATGTGAATTTCCAGTGGCCATCAAAGTGTGAGGAAACACATCC 2280  
 QY 653 ----- 653  
 Db 2281 CCCAAAGCCAAACAGAAATCTTAGCAGAGCATAGCTGATGGTGGTGGGCTCCCA 2340  
 QY 653 ----- 653  
 Db 2341 TATGTCTCCGCTTCTGGGCATCTGCCCTGACATCCACGGTGCAGCTGTGACACAGCTT 2400  
 QY 653 ----- 653  
 Db 2401 ATGCCCTATGGTGCCTCTTAGACCATGTCCGGGAAACCGGAGCGCTGGGCTCCCAG 2460  
 QY 653 ----- 653  
 Db 2461 GACCTCTGAAGTGTGTATGAGATGTCAGATGCCAAGGGATGAGTACTGGAGGATGGCG 2520  
 QY 653 ----- 653  
 Db 2521 CTCGTACACAGGACTTGGCCGCTCGAAGCTGCTGGTCAAGAGTCCCAACCATGTCAA 2580  
 QY 653 ----- 653  
 Db 2581 ATTACAGACTTCGGGCTGGCTGGCTGGACATTGACGAGACAGATACCATTGCAGAT 2640  
 QY 653 ----- 653  
 Db 2641 GGGGGCAAGGTGCCCATCAAGTGTGGTGGCGTGGAGTCCATTCTCGCGCGGGTTACAC 2700  
 QY 653 ----- 653  
 Db 2701 CACCAGATGATGTGGAGTTATGGTGTACTGTCTGGAGCTGATGACTTTTGGGGCC 2760  
 QY 653 ----- 653

Db 2761 AAACCTTAGCATGGATCCACCGCGGAGATCCCTGACCTGCTGGAAGAGGGGAGCGG 2820  
 QY 653 ----- 653  
 Db 2821 CTGCCCCAGCCCCCATCTGCACCATGTGATGTCTATCATGATCATGTCAATGTTGGATG 2880  
 QY 653 ----- 653  
 Db 2881 ATTGACTCTGAATGTGCGGCCAAGATTCCGGGAGTTGGTCTCTGAATTCTCCCGCATGGCC 2940  
 QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664  
 Db 2941 AGGGAGCCCCAGCGCTTTGTGTGTCATCCAGAAATGAGGACTTGGGCCCAACCCAGTCCCTTG 3000  
 QY 665 AspSerThrPheTyrArgSerLeuGluAspAspMetGlyAspLeuValAspAla 584  
 Db 3001 GACAGACCTTCTACCGCTCACTGTGTGGAGACGATGACATGGGACCTGGTGGATGCT 3060  
 QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704  
 Db 3061 GAGGAGTATCTGGTACCCAGCAGGGCTTCTTCTGTCCAGACCTGCCCGGGCGCTGGG 3120  
 QY 705 GlyMetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyGlyAspLeuThr 724  
 Db 3121 GGCATGGTCCACACAGCGCAGCTCATCTACAGGAGTGGCGGTGGGACCTGACAC 3180  
 QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744  
 Db 3181 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCCAGGTCTCCACTGCGACCTCCGAAGGG 3240  
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764  
 Db 3241 GCTGGCTCGATGATTTGATGGTACCTGGGAATGGGGCGACGCCAAGGGCTGCAAGC 3300  
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
 Db 3301 CTCCACACATAGACCCACGCCCTCTACAGCGGTACAGTGGAGACCCACAGTACCCCTG 3360  
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
 Db 3361 CCCTCTGAGACTGATGGCTACGTGGCCCTTGCCTGACCTGCAGCCCCCAGCTGAATATG 3420  
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824  
 Db 3421 AACCGACCATGATGTCGGCCCCAGCCCCCTTCGCCCGAGAGGGCCCTCTGCTGCTGCC 3480  
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
 Db 3481 CGACCTGCTGTGCCACTCTGGAAGGGCCAAAGACTCTCTCCCCAGGGGAAGATGGGGTC 3540  
 QY 845 ValLysAspValPheAlaPheGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
 Db 3541 GTCAAGAGCTTTTGGCTTTGGGTGCGCGTGGAGAACCCGAGTACTTGACACCCAG 3600  
 QY 865 GlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeu 884  
 Db 3601 GGAGGAGCTGCCCTCAGCCCCACCCCTCTCTGCTTCCAGCCAGCCTTCGACACACCTC 3660  
 QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
 Db 3661 TATTACTGGGACGAGCACCACAGAGCGGGGGCTCCACCCAGCAGCCTTCAAAGGAGCA 3720  
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
 Db 3721 CCTACGGCAGAAACCCAGAGTACCTGGGTCTGGACGTGCCAGTG 3765

RESULT 12  
 AR202597  
 LOCUS  
 DEFINITION  
 AR202597  
 ACCESSION  
 VERSION

AR202597 4530 bp DNA linear PAT 20-JUN-2002  
 Sequence 4 from patent US 6365151.  
 AR202597  
 AR202597.1 GI:21498768

KEYWORDS

SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 4530)

AUTHORS Halpern, M.S. and England, J.M.

TITLE Cellular immunogens comprising cognate proto-oncogenes

JOURNAL Patent: US 6365151-A 4 02-APR-2002;

FEATURES

Location/Qualifiers

1..4530

source

BASE COUNT 922 a 1382 c 1346 g 880 t

ORIGIN

Alignment Scores:

Pred. No.: 1.34e-186 Length: 4530

Score: 4892.00 Matches: 918

Percent Similarity: 73.15% Conservative: 0

Best Local Similarity: 73.15% Mismatches: 1

Query Match: 96.34% Indels: 336

DB: 6 Gaps: 1

SE06 (1-919) x AR202597 (1-4530)

Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20  
Db 151 ATGGAGCTGGCGGCTTGTGCGGCTGGGGCTCTCTCGCCCTTTGCCCGCGAGCC 210  
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
Db 211 GCGAGCACCCAAAGTGTGCACGGCACACATGATGAAGCTGGCGCTCCCTGCCAGTCCCGAG 270  
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrglnGlyCysGlnValGlnGlyAsnLeu 60  
Db 271 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGGAAACCTG 330  
Qy 61 GluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
Db 331 GAACCTACCTACCTGCCACCAATGCCAGCTGTCTCTTCTGCAGGATATCCAGGAGGTG 390  
Qy 81 GlnGlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
Db 391 CAGGGCTACGTGCTCATCGCTCACCAAGTAGGAGGAGTCCCACTGCAGAGGCTCGG 450  
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAspAsnGly 120  
Db 451 ATGTGGCAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGGTGCTAGACAAATGGA 510  
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
Db 511 GACCCGTGAACAATACCAACCCCTGTACAGGGGCTTCCCGAGGAGGCTCGGAGCTG 570  
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160  
Db 571 CAGCTTCCAGACCTTCACAGAGATCTTGAAGGAGGGTCTTGATCCAGCGGAACCCCGAG 630  
Qy 161 LeuCysTyrrGlnAspThrIleLeuTyrrLysAspIlePheHisLysAsnAsnGlnLeuAla 180  
Db 631 CTCTGCTACCAAGACAGATTTGTGAAGGACATCTTCCACAGAACCAACAGCTGGCT 690  
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
Db 691 CTCACACTGATAGACACCAACCGCTCTCGGCGCTGCCACCCCTGTCTTCTCCGATGTGAAG 750  
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
Db 751 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTGAGAGCTTCAGAGCCCTGAGCGCACTGTCTGT 810  
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240  
Db 811 GCCGGTGGCTGTGCCCTGCAAGGGCCACTGCCCCACTGACTGCTGCCATGAGCAGTGT 870  
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260

Db 871 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCCTGGCCTCCCTCCACTTCAACAC 930  
Qy 261 SerGlyTleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThrPheGlu 280  
Db 931 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGTGTACCTTACACACAGACACGCTTTGAG 990  
Qy 281 SerMetProAsnProGluGlyArgTyrrThrPheGlyAlaSerCysValThrAlaCysPro 300  
Db 991 TCCATGCCCAATCCGAGGCGCGGTATACATTCGCGCCAGCTGTGTGACTGCTGTCTCC 1050  
Qy 301 TyrAsnTyrrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
Db 1051 TACAACATACCTTTCTAGCGACGTGGATCTGTGCACCTCTGCTGCCCTGCACACGAA 1110  
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
Db 1111 GAGGTGACAGCAGAGGATGAACACAGCGGTGTGAGAAGTGCACAGACCCCTGTGCCGA 1170  
Qy 341 ValCysTyrrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
Db 1171 GTGTGCTATGCTGGGCATGGACACTTGGAGAGGTGAGGGCAGTTACCACTGCCAAT 1230  
Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
Db 1231 ATCCAGGAGTTTGTGCTGCAAGAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC 1290  
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
Db 1291 TTTGATGGGACCCAGCCTCCAACTGCCCCGCTCCAGCCAGAGAGCTCCAAGTCTTT 1350  
Qy 401 GluThrLeuGluGluIleThrGlyTyrrLeuTyrrIleSerAlaTrpProAspSerLeuPro 420  
Db 1351 GAGACTCTGGAAGAGATCAGAGTTACTATACATCTCAGCATGCCGCGGACAGCCTGCCT 1410  
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
Db 1411 GACCTCAGCTCTTCCAGAACCTGCAGAGTAATCGGGGAGCAATTTCTGCAATATGGGCC 1470  
Qy 441 TyrrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
Db 1471 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCTCACTCAGGAGAA 1530  
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480  
Db 1531 CTGGCAGTGGAGTGGCCCTCATCCACCATTAACACCCCTCTGCTTCTGTCGACACCGGTG 1590  
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
Db 1591 CCCTGGGACCACTCTTTCCGAAACCCGACCAAGCTCTGTCTCCACACTGCAACCGGCCA 1650  
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
Db 1651 GAGGACAGAGTGTGGCGGAGGGCTGGCCCTGCCACAGCTGTGCGCGGAGGCACTGC 1710  
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
Db 1711 TGGGTCACAGGGCCACCCAGTGTCTCACTGACGACAGTTCTTCTCGGGCCAGGAGTGC 1770  
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrrValAsnAlaArgHisCys 560  
Db 1771 GTGGAGGAATGCCAGTACTGTCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1830  
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
Db 1831 TTCCCGTGGCCACCTGAGTGTGACGCCCCAGAAATGGCTCAGTACCTGTTTGTGACCGGAG 1890  
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrrLysAspProPheCysValAlaArgCys 600  
Db 1891 GCTGACAGTGTGGCTGTGCCCTGCTCCACTATGAAGACCTCCCTCTCTGCGTGGCCCGCTGC 1950  
Qy 601 ProSerGlyValLysProAspLeuSerTyrrMetProIleTrpLysPheProAspGluGlu 620

Db 1951 CCCAGGGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG 2010  
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640  
 Db 2011 GGGGATGCCAGCTTGGCCATCAACTGACCCACTCTCTGTGGACTGGATGACAG 2070  
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653  
 Db 2071 GGCTGCCCGCGGAGCAGAGAGCCAGCCCTCTGACGTCCATCGTCTCTGCGGTGTGGC 2130  
 QY 653 ----- 653  
 Db 2131 ATTCTGCTGGTGGTCTTGGGGTGGTCTTTTGGGATCTCTCATCAAGCAGCGGAGCAG 2190  
 QY 653 ----- 653  
 Db 2191 AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAACGGAGCTGGTGGAGCGGTG 2250  
 QY 653 ----- 653  
 Db 2251 ACACCTAGCGGAGCGATGCCAACACGCGCAGATGCGGATCCTGAAAGAGACGAGCTG 2310  
 QY 653 ----- 653  
 Db 2311 AGGAAGGTGAAGGTGCTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2370  
 QY 653 ----- 653  
 Db 2371 CCTGATGGGAGAATGTGAATTCAGTGGGCCATCAAGTGTGGGGAACACATCC 2430  
 QY 653 ----- 653  
 Db 2431 CCCAAAGCCAAACAAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGGCTCCCA 2490  
 QY 653 ----- 653  
 Db 2491 TATGTCTCCCGCTTCTGGGCACTGCGCTGACATCCACGGTGCAGCTGGTGACAGCTT 2550  
 QY 653 ----- 653  
 Db 2551 ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCGGAGCGCTGGGCTCCAG 2610  
 QY 653 ----- 653  
 Db 2611 GACCTGCTGAAGTGTGTATGAGATGTGCAAGGGGATGAGTACCTGGAGGATGTGGG 2670  
 QY 653 ----- 653  
 Db 2671 CTCGTACACAGGACTTGGCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAA 2730  
 QY 653 ----- 653  
 Db 2731 ATTACAGACTTCGGCTGGCTGGCTGCTGGACATTGAGGACAGAGTACCATCCAGAT 2790  
 QY 653 ----- 653  
 Db 2791 GGGGCAAGGTGCCCATCAAGTGGATGGCGTGGAGTCCATTCTCCGCGGGGTTCACC 2850  
 QY 653 ----- 653  
 Db 2851 CACCAGATGATGTGTGGATTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGCGC 2910  
 QY 653 ----- 653  
 Db 2911 AAACCTTAGTGGATCCACGCCGGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2970  
 QY 653 ----- 653  
 Db 2971 CTGCCCCAGCCCCCATCTGCACCATGTGATGCTACATGATGATGCTCAATGTTGGATG 3030  
 QY 653 ----- 653  
 Db 3031 ATTGACTCTGAATGTCCGCCAAGATTCCGGGAGTTGGTGTCTGAATCTCCCGCATGCC 3090

QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664  
 Db 3091 AGGAGCCCCCAGCGCTTGTGTCTATCCAGATGAGGACTTGGGCCACCCAGTCCCTTG 3150  
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684  
 Db 3151 GACAGCACCTTCTACCGCTCACTGTGGAGGACGATGACATGGGGACCTGGTGGATGCT 3210  
 QY 685 GluGluTyrLeuValProGlnGlnPhePheCysProAspProAlaProGlyAlaGly 704  
 Db 3211 GAGGAGTATCTGTACCCAGCAGGCTTCTTCTGTCCAGACCTGCCCGGGCGGTGG 3270  
 QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724  
 Db 3271 GCATGGTCCACACAGGACCGCAGCTCATCTACAGAGGTGGGGTGGGACCTGACA 3330  
 QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744  
 Db 3331 CTAGGCTGGAGCCCTCTGAAGAGGAGGCCCCAGCTCTCCACTGSCACCTCCGAAGG 3390  
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764  
 Db 3391 GCTGGCTCCGATGATATTTGATGGTACCTGGGAATGGGGGAGCCCAAGGGGCTGCAAGC 3450  
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
 Db 3451 CTCGCCACACATGACCCCGCTCTACAGCGGTACAGTGGAGACCCACACACTACCCCTG 3510  
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
 Db 3511 CCCTCTGAGACTGATGGCTACGTTGCCCTTACCTGACGCCCGCAGCTGAATATGTC 3570  
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824  
 Db 3571 AACCAAGCAGATGTTGGCGCCCGCAGCCCTTCCGCCCGAGAGGGCCCTTGCTGCTGCC 3630  
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
 Db 3631 CGACCTGTGGTGGCCACTCTGGAAGGGCCAGACTCTCTCCCGAGGAGAAATGGGTC 3690  
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
 Db 3691 GTCAAAGACGTTTTTGGCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCAG 3750  
 QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884  
 Db 3751 GGAGAGGTGGCCCTCAGCCCCACCTCTCTCTGCTTTCAGCCCGCCTTCGACACACCTC 3810  
 QY 885 TyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
 Db 3811 TATTACTGGGACGAGACCCAGAGGGGGGCTCCACCCAGCAGCCTTCAAGGGGACA 3870  
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
 Db 3871 CCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGAGCTGCCAGTG 3915

RESULT 13  
 AR283481  
 LOCUS  
 DEFINITION  
 AR283481  
 ACCESSION  
 VERSION  
 AR283481.1  
 GI:29720342  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Unknown.  
 Unclassified.  
 1 (bases 1 to 4530)  
 Nicorette,C.A.  
 Therapeutic compounds  
 Patent: US 6528060-A 1 04-MAR-2003;  
 Location/Qualifiers  
 1..4530  
 source

AR283481  
 Sequence 1 from patent US 6528060.  
 linear PAT 10-APR-2003

```

BASE COUNT      922 a 1382 c 1346 g      880 t
ORIGIN
Alignment Scores:      1.34e-186      Length:      4530
Pred. No.:      4892.00      Matches:      918
Score:      73.15%      Conservative:      0
Percent Similarity:      73.15%      Mismatches:      1
Best Local Similarity:      96.34%      Indels:      336
Query Match:      6      Gaps:      1
DB:
SEQ6 (1-919) x AR283481 (1-4530)
Qy      1      MetGluLeuAlaLaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGluVala 20
Db      151      ATGGAGCTGGGGCGCTGTGGCGCTGGGGGCTCTCTCGGCCCTTGGCCCCCGGAGCC 210
Qy      21      AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db      211      GCGAGCACCCAAAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 270
Qy      41      ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnLeu 60
Db      271      ACCCACCCTGGACATGCTCCCGCCACCTCTTACCAGGGCTGCCAGGTGGTGCAAGGAAACCTG 330
Qy      61      GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db      331      GAATCACCCTACCTGCCACCAATGCCAGCTGTCCTTCCTGCGAGGATATCCAGGAGGTG 390
Qy      81      GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      391      CAGGGCTACGTGCTCATCGCTCACAAACCAAGTAGGAGGCTGCCACCTGCAGAGGCTGCGG 450
Qy      101      IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      451      ATTGTGGGAGGACCCAGCTCTTTGAGGACAACTATGCCCTTGGCCGTGCTAGACAAATGGA 510
Qy      121      AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db      511      GACCCGTGAACATACCAACCCCTGTACAGGGGCTTCCCAGAGGGCTTGGGGAGCTG 570
Qy      141      GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db      571      CAGCTTGAAGCCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCCAG 630
Qy      161      LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db      631      CTCTGCTACCAAGGACAGATTTGTGGAAGGACATCTTCCACAAAGAACCAACCACTGGCT 690
Qy      181      LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      691      CTCACACTGATAGACACCAACCCGCTCTCGGGCTTCCACCCCTGTTCTCCGATGTGTAA 750
Qy      201      GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      751      GGCTCCCGCTGCTGGGAGAGAGTCTCAGAGATTGTTCAGAGCCCTGACGCGCACGTCTGT 810
Qy      221      AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db      811      GCCGGTGGCTGTGCCCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 870
Qy      241      AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db      871      GCTGCCGGCTGCAGGGGCCCAACCACTCTGACTGCTGCTGGCTGCCCTCCACTTCAACAC 930
Qy      261      SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db      931      AGTGGCATCTGTAGCTGCACCTGCCAGCCCTGCTACCTTACAAACACACAGACACCTTTGAG 990
Qy      281      SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db      991      TCCATGCCCAATCCCGAGGGCCGGTATACATTCGGGCCAGCTGTGTGACTGCTGCTGCC 1050
```

```

Qy      301      TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db      1051      TACAACCTACCTTTCTACGGAGCTGGGATCTTCGACCCCTCGTCTGCCCCCTGCACAAACCA 1110
Qy      321      GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db      1111      GAGCTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAAGTCAGCAAGCCCTGTGCCCGA 1170
Qy      341      ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db      1171      GTGTGCTATGCTTGGGCATGGAGCACTTGGAGAGGTGAGGGCAGTTTACCAGTGCCTAAT 1230
Qy      361      IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db      1231      ATCCAGGAGTTGCTGGCTGCAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGC 1290
Qy      381      PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db      1291      TTTGATGGGACCCAGGCTCCAAACACTGCCCCGCTCCAGCCAGCAGCAGCTCCAAGTGT 1350
Qy      401      GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db      1351      GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGGCCGAGCAGCTGCCT 1410
Qy      421      AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db      1411      GACCTCAGCGTCTTCCAGAACCTGCAAGTAATATCGGGGACCAANTTCTGCACAAATGGCC 1470
Qy      441      TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db      1471      TACTCGCTGACCCGTGAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGAA 1530
Qy      461      LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db      1531      CTGGGCACTGAGTGGGCTTGGCCCTCATCCACATAACACCCACCTGCTGCTGCACACG 1590
Qy      481      ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db      1591      CCCTGGGACCACTCTTTCCGAACCCGACCAAGCTGCTGCTCCACACTGCCCAACCGGCA 1650
Qy      501      GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db      1651      GAGGACGAGTGTGTGGCGAGGGCTGGCCCTGCCACCACTGCTGCCCGGAGGACCTGC 1710
Qy      521      TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db      1711      TGGGGTCCAGGGCCCACTGCTGCTCAACTGCAGCAGTTCTTTCGGGGCCAGGAGTGC 1770
Qy      541      ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db      1771      GTGGAGGAATGCCAGGTACTGCAGGGCTCCCCAGGAGGATATGTGAATGCCAGGCACTGT 1830
Qy      561      LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db      1831      TTGCGCTGCCACCTTGTAGTGTGACGCCAGAAATGGCTCAGTGACCTGTTTTGGACCGAG 1890
Qy      581      AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db      1891      GCTGACCACTGTGGGCTGTGCCCACTATAAGGACCCCTCCCTTCTGCGTGGCCCGCTGC 1950
Qy      601      ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db      1951      CCCAGCGGTGTGAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010
Qy      621      GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db      2011      GGGCATGCGGACCTTGGCCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 2070
Qy      641      GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
Db      2071      GGCTGCCCCCGGAGCAGAGAGCCGCTCTGACGTCCATCGCTCTCTGCGGTGGTGGC 2130
```





/translation="MELAALCRWGLIALLPAGAASTQVCTGCTDMKRLRLPASPEETHLD  
MLRHLQVQGVGNLELYLPTNASISFLDQIEQVGYVYLIAHNVQVROLRLTV  
RQQLQEDFNVALVLDNGDPLNNTTPTWGPASGGLRELQRLSLEILKGVILQRNPQ  
LCQDTILMRDIFHKNNQALATLIDTNRSRACHPCSPMKAGSRGWESSDDCSLRT  
VCAGGARCNGPLPTDCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNT  
DTPESMPPEGRYTFGASCCTACGCTNYLSDTGVCTLCVCPHNOEVAEDGTQRCCK  
SKPCARVCYGLGMEHLEVRVTSANTQEPAGCKIFGSLAFLPESFGDGPASNTAPL  
QPQLOVFEETLEETGYLISAWPDSLPDLVSFQNLQVIRGRILHNGAYSLTLQGLI  
SWLGLSLRHGSGSLAIHNTHLCLFVHTVPMDLFRNPHOALLHTRANRPEDECYEG  
LACHQLCARHGNGPCTQVCNCSQFLRGQCEVEECRVLQGLPREYVNAHRLCLPHE  
CQONGSVTCFGEADQCVACAHYKDPFCVACRCPGVKPDLSLTPYIWRPFDEEACQ  
PCINTHSCVDLDDGCPAEQASPLSTSVASVGLLVGLVGVFGILIKRRQKI  
RKYTMRLQETDELVEPLTPSGAMPNOAMRILKETELRKVKVLGSGAFGTVYKGLWI  
PDGENKPIVAIKRENTSPKANKETLDEAYVMAGVGSYPVSRLLGICLSTVQLVT  
OLMPYGLDLHVRNRRGLSQDLNWCMIKAGMSYLEDVRLVRLHDLAARNVLVSP  
NHVKITDFGLARLLDDEYHADGGKVPKIMWALRESILLRRTHOSDWSYGVTVWE  
LMTFGAKDPQIFAREIPDLLEGERLPPQPICTIDYIMIMVKCWDISCRPRFREL  
VSEFSRNARDQRFVYQIQLDGPASPLDSTFYRSLEDDMDGDLDAEYLYPQQGF  
FCDDPAPGAGMYHHRSSSTRSGGDLTLGLEPSEEEAPRSLAPSEAGSDVFDG  
DLGMGAAGLQSLTHDPSPLQRYSEDPTVPLPSETDGVVAPLTCSPQPEYVNPQDVR  
POPPSPREGPLPAARAGATLERAKTLPCKNGVVKDVFAGGAVENPEYLTTPQGGAA  
POPHPPAFSPADNLYYWDODPPERGAPSTFKGTPYTAENPEYVLGLDYPV"

BASE COUNT 922 a 1382 c 1346 g 880 t  
ORIGIN

## Alignment Scores:

Pred. No.: 1.34e-186 Length: 4530  
Score: 4892.00 Matches: 918  
Percent Similarity: 73.15% Conservations: 0  
Best Local Similarity: 73.15% Mismatches: 1  
Query Match: 96.34% Indels: 336  
DB: 6 Gaps: 1

SEQ6 (1-919) x AX282577 (1-4530)

Qy	1	MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla	20
Db	151	ATGGAGTGGCGGCTTGTCCGCTGGGGCTCCTCTCCCTCCCTTGTCCGCCCGGAGCC	210
Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	211	GCAGACCCCAAGTGTGCACCGGCACAGACATGAAGTGGCGCTCCCTGCCAGTCCCGAG	270
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlyAsnLeu	60
Db	271	ACCCACCTGGACATGCTCCGCCACCTCTACACAGGCTGCCAGTGGTGCAGGGAACCTG	330
Qy	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	331	GAACCTACCTACCTGGCCCAATGCCAGCCTGTCTCTTCCCTGCCAGGATATCCAGGAGGTG	390
Qy	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	391	CAGGCTACCTGTCTATCGCTACAAACAAAGTGGGAGGTGCCACTGCCAGAGGCTGCGG	450
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	451	ATTGTGGAGCACCCAGCTCTTTGAGGACAACTATGCTTGGCCGTGCTAGACATGGA	510
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyLeuArgGluLeu	140
Db	511	GACCCGCTGAACAATACACCCCTGTCCAGGGGCCCTCCAGGAGGCTGCGGAGCTG	570
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160
Db	571	CAGCTCGAAGCTTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCCAG	630
Qy	161	LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
Db	631	CTCTGCTACCCAGGACAGATTTTGTGAAGAGACATCTTCACAGAAACAACACAGCTGGCT	690
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db			
691	CTCACACTGATAGACACCAACCGCTCTCGGGCCCTGCACCCCTCTTCTCCGATGTGTAAG	750	
Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	751	GGCTCCCGCTGTCTGGGAGAGAGTCTTGAGGATTTGCAGAGCCTGCAGCGCACCTGTCTGT	810
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	811	GCCGTTGGCTGTGCCCTGCCAGGGGCCACTGCCACTGACTGCTGCCCTGCCTCAACAC	870
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	871	GCTGCCGCTGCACGGGCCCAAGACACTCTGACTGCTGCCCTGCCCTGCCTCAACAC	930
Qy	261	SerClyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	931	AGTGGCATCTGTGAGCTGCACTGCCACGCCCTGTGTACCTACAAACACAGACACACTTTGAG	990
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	991	TCCATGCCCAATCCGAGGGCCGATATACATTGCGGCCAGCTGTGTGACTGCCTGTCCC	1050
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	1051	TACAACTACCTTTCTACGGACGTGGGATCCTGCACCCCTCGCTGCCCTGCACACCA	1110
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	1111	GAGGTGACACAGAGGATGAACACACGCGGTGTGAGAAGTGCACAGACCCCTGTGCCCGA	1170
Qy	341	ValCystTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1171	GTGTGCTATGCTTGGCATGGAGCACTTCCGAGAGGTGAGGCGAGTACCAGTGCCTCAAT	1230
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1231	ATCCAGAGTGTGCTGGCTGCAAGAGATCTTTGGGAGCCTGGCATTTCTGCCGAGAGC	1290
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnGlnValPhe	400
Db	1291	TTTGTATGGGACCCAGCCTCCACACTGCCCGCTCCAGCGCAGCAGAGCTCCAAGTGT	1350
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1351	GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGACAGCCTGCCT	1410
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1411	GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGGAGCAATTCGCAATGGGCCC	1470
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1471	TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGGCTCCTACTGAGGAA	1530
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1531	CTGGCAGCTGACTGGCCCTCATCCACCATACACCCACCTCTCTCTTCGTGCACACCGTG	1590
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1591	CCCTGGACACGAGCTCTTTCGGAACCCCGCACCAAGCTCTGTCTCCACACTGCCAACCGCCA	1650
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1651	GAGGACAGTGTGTGGGCGAGGGCTTGGCTGCCCTGCCAGCTGTGGCGCCGAGGCACTGC	1710
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1711	TGGGTCACGGCCACCCAGTGTGTCACTGCAGCAGCTTCTCTTCGGGGCCAGGAGTGC	1770
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1771	GTGGAGGAATGCCAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1830

QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 DB 1831 TTGCGGTGCCACCTGAGTGTGAGCCCAAGATGCTCAGTACCTGTTTGGACCGGAG 1890  
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
 DB 1891 GGTGACAGTGTGTGGCTGTGCCACTATAAGGACCTCCCTTCTGCGTGGCCGCTGC 1950  
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTyrPheProAspGluGlu 620  
 DB 1951 CCAGCGGTGTGAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010  
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640  
 DB 2011 GCGCATGCCAGCTTGCCCACTCAACTGCACCCACTCTCTGTGTGACCTGGATGACAG 2070  
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653  
 DB 2071 GGCTGCCCGCGGAGCAGAGAGCCAGCCCTCTGAGGTCCATCTCTGCGGTGGTTGGC 2130  
 QY 653 ----- 653  
 DB 2131 ATTCTGCTGGTGTGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCGAGCGGAGCAG 2190  
 QY 653 ----- 653  
 DB 2191 AAGATCCGGAAGTACACGATGCGAGAGTGTCTGCAGAAACGAGCTGTGTGAGCGCGTG 2250  
 QY 653 ----- 653  
 DB 2251 ACACCTAGCGGAGCGATGCCCAACCAAGCGCGAGATGCGGATCCTGAAAGAGACGAGCTG 2310  
 QY 653 ----- 653  
 DB 2311 AGGAAGTGAAGTGTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2370  
 QY 653 ----- 653  
 DB 2371 COTGATGGGAGAAATGTGAATTCAGTGCCCATCAAAGTGTGAGGAAACACATCC 2430  
 QY 653 ----- 653  
 DB 2431 CCCAAAGCAACAAGAAATCTTAGACAGAGCATACGTGATGGCTGTGTGGGCTCCCA 2490  
 QY 653 ----- 653  
 DB 2491 TATGTCCTCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGTGTGTGACACAGCTT 2550  
 QY 653 ----- 653  
 DB 2551 ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCGCGAGCGCTGGGCTCCCA 2610  
 QY 653 ----- 653  
 DB 2611 GACCTGCTGAAGTGTGTATGAGATGTCGAAGGGATGAGTACCTGGAGGATGTGCGG 2670  
 QY 653 ----- 653  
 DB 2671 CTCGTACAGGGACTTGGCCSCTCGGAACGTGTGGTCAAGAGTCCCAACCATGTCAA 2730  
 QY 653 ----- 653  
 DB 2731 ATTACAGACTTCGGGTGCTCGGCTGTGGATTTAGGACATTGAGGACAGAGTACCATGCAGAT 2790  
 QY 653 ----- 653  
 DB 2791 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGTTTCCAC 2850  
 QY 653 ----- 653  
 DB 2851 CACCAGATGATGTGGAGATTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2910

QY 653 ----- 653  
 DB 2911 AAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGCTGAAAGGGGAGCGG 2970  
 QY 653 ----- 653  
 DB 2971 CTGCCCCAGCCCCCATCTGSCACCATTTGATGTCTACATGATCATGTCCTCAATGTTGGATG 3030  
 QY 653 ----- 653  
 DB 3031 ATTGACTCTGAATGTGGCCCAAGATTCCGGGAGTTGGTCTCTGAATTCTCCCGCATGGCC 3090  
 QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664  
 DB 3091 AGGAGCCCCACGCGCTTTGTGTCTATCCAGATAGGACTTGGGCCACCCAGTCCCTTG 3150  
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684  
 DB 3151 GACAGCACCTTCTACCGCTCACTGTGTGGAGGACGATGACATGGGACCTGGTGGATGCT 3210  
 QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704  
 DB 3211 GAGGAGTATCTGTATCCCGCAGCGGCTTCTTCTGTCCAGACCTGCCCGGGCGCTGG 3270  
 QY 705 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 724  
 DB 3271 GGCATGGTCCACACAGGACCGCAGCTCATCTACAGGAGTGGCGTGGGACCTGACA 3330  
 QY 725 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 744  
 DB 3331 CTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCGCCAGGTCTCCACTGGCACTCCCGAAGGG 3390  
 QY 745 AlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSer 764  
 DB 3391 GCTGGCTCGATGATTTGATGGTACCTGGGAATGGGGCGACCCAGGGGTGCCAAGC 3450  
 QY 765 LeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeu 784  
 DB 3451 CTCCCCACACATGACCCCGCCCTCTACAGCGGTACAGTGGAGACCCACAGTACCCCTG 3510  
 QY 785 ProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrVal 804  
 DB 3511 CCCTCTGAGACTGATGGCTACGTTGCCCTTGCCTGCAGCCCGCCAGCTGAATATGTG 3570  
 QY 805 AsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAla 824  
 DB 3571 AACAGCCAGATGTTTGGCCCCAGCCCCCTTCCGCCCGAGAGGGCCCTCTGCTGCTGCC 3630  
 QY 825 ArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGlyVal 844  
 DB 3631 CGACCTGTGTGTCACCTCTGGAAAGGGCCAAAGACTCTCTCCCGAGGAGAAATGGGTC 3690  
 QY 845 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGln 864  
 DB 3691 GTCAAAGAGTGTGTTTGGCTTGGGGGTGCCGGAAGACCCGAGTACTTGACACCCAG 3750  
 QY 865 GlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsnLeu 884  
 DB 3751 GGAGGAGTGGCCCTCAGCCCCACCTCTCTCTGCTTCCAGCCAGCTTCCGACACCTC 3810  
 QY 885 TyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThr 904  
 DB 3811 TATTACTGGGACGAGCCACAGAGCGGGGGTCCACCCAGGACCTTCAAAGGAGACA 3870  
 QY 905 ProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919  
 DB 3871 CCTACGGCAGAACCCAGAGTACCTGGGTCTGGAGTGGCCAGTG 3915  
 RESULT 15  
 AX587649  
 LOCUS AX587649 4530 bp DNA linear PAT 10-JAN-2003  
 DEFINITION Sequence 119 from Patent WO0246467.  
 ACCESSION AX587649

```

VERSTON AX587649.1 GI:28212358
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE artificial sequences.
1
AUTHORS Bertucci,F., Houlgatte,R., Birnbaum,D., Nguyen,C., Viens,P. and
Fert,V.
TITLE Gene expression profiling of primary breast carcinomas using arrays
of candidate genes
JOURNAL Patent: WO 0246467-A 119 13-JUN-2002;
FEATURES Ipsogen (FR)
source Location/Qualifiers
1..4530
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="primer"
misc_feature 1..4530
/note="v-erb-b2 avian erythroblastic leukemia viral
oncogene homolog 2 (neuro/glioblastoma derived oncogene
homolog) (ERBB2) gene."
BASE COUNT 922 a 1382 c 1346 g 880 t
ORIGIN

Alignment Scores:
Pred. No.: 1.34e-186 Length: 4530
Score: 4892.00 Matches: 918
Percent Similarity: 73.15% Conservative: 0
Best Local Similarity: 73.15% Mismatches: 1
Query Match: 96.34% Indels: 336
DB: 6 Gaps: 1

SEQ6 (1-919) x AX587649 (1-4530)
Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProProGlyAla 20
Db 151 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCTCTCCGCGCTCTTCCGCGCGGAGCC 210
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 CGGAGCACCAAGTGTCCCGGCACAGACATGAAGCTGGCGCTCCCTGCCAGTCCCGAG 270
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAACCTG 330
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 331 GAACCTACCTACCTGCCCGCAATATGCCAGGCTGTCTTCCTGCCAGGATATCCAGGAGGTG 390
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTACGTGTCTATCGCTACAAACAAAGTAGGAGGTGCCACCTGCAGAGGCTCGG 450
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATGTGTGGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGGA 510
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGTGAACAATACCAACCCCTGTACAGGGGCCCTCCCGAGGAGGCTCGCGGAGCTG 570
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 571 CAGCTTCGAAGCCCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCGAG 630
Qy 161 LeuCystTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGCTACCGAGGACGATTTGTGAAGGAGCATCTTCACAGAGAACACCACTGGCT 690
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACATGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAG 750

```

```

201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
|
|
|
751 GGCTCCCGCTGCTGGGAGAGAGATTCTGAGGATTGTTCAGAGCCTGAGCGGCTGTCTGT 810
|
|
|
221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
|
|
|
811 GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCACTGACTGCTGCTCATGAGCAGTGT 870
|
|
|
241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
|
|
|
871 GCTGCCGGCTGCAGGGGCCCAAGCACTCTGACTGCTGGCTGCTGCCCTCCACTTCAACAC 930
|
|
|
261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
|
|
|
931 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGTACCTACACACAGACACCTTTGAG 990
|
|
|
281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
|
|
|
991 TCCATGCCCAATCCGAGGGCGGTATACATTCCGGCGCAGCTGTGTACTGCTGCTGCC 1050
|
|
|
301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
|
|
|
1051 TACAACCTACCTTTCTACGGAGCTGGGATCTCTGCACCTCGTCTGCCCTCCCTGCACAA 1110
|
|
|
321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
|
|
|
1111 GAGGTGACAGCAGAGGATGAACACACAGCGGTGTGAGAAGTGACAGACCCCTGTGCCGA 1170
|
|
|
341 ValCystTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
|
|
|
1171 GTGTGCTATGCTTGGGCATGGAGCACTTGGAGAGGTGGAGGCGAGTTACCAATGCCAAT 1230
|
|
|
361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
|
|
|
1231 ATCCAGAGATTGCTGGCTGCAAGAAGATCTTTGGAGCCTTGGCATTTCTCCCGGAGAGC 1290
|
|
|
381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
|
|
|
1291 TTTGATGGGAGCCAGCCCTCAACACGTGCCCGCTCCAGCCAGCAGAGCAGCTCCAAGTGT 1350
|
|
|
401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
|
|
|
1351 GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGCCGCGGACAGCTGCCT 1410
|
|
|
421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
|
|
|
1411 GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGAGCGAATTTCTGCAAAATGGCGCC 1470
|
|
|
441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
|
|
|
1471 TACTCGCTGACCTTGCAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAA 1530
|
|
|
461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
|
|
|
1531 CTGGGCACTGGAGCTGGCGCTCATCCACCAATAACACCCACTCTCTGCTGCTGCACACGGTG 1590
|
|
|
481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
|
|
|
1591 CCCTGGACCAAGCTCTTTTCGGAACCCGACCAAGCTCTGCTGCCACTGCTGCCAACCGGCCA 1650
|
|
|
501 GluAspIleCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
|
|
|
1651 GAGGACAGGTGTGTGGCGAGGGCTGGCTGCCACCAAGCTGTGCTGCCCGCGGAGGCACTGC 1710
|
|
|
521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
|
|
|
1711 TGGGGTCCAGGGCCCAACCCAGTGTGTCACTGACAGCAGTCTCTTCGGGGCCAGGAGTGC 1770
|
|
|
541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
|
|
|
1771 GTGAGGAATGCCAGTACTGCGAGGGCTCTCCCGAGGAGTATGTGAATGATGCCAGCACTGT 1830

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 461.761 Seconds  
(without alignments)  
3105.423 Million cell updates/sec

Title: SEQ5  
 Perfect score: 326  
 Sequence: 1 ONEDLGASPPLDSTFYRSLT.....GFFCPDPAPGGGVHHRHR 59

Scoring table: BLQSUN62

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

MAXIMUM OF 20000000

Post-processing: Minimum Match 0%

Maximum Match 100%

## Listing first 45 summaries

Command line parameters:

```

Command line parameters:
-MODEL=framer*_p2n.model -DEV=xlp
-C/cgna2/USPTO_spool_P/HOLLERAN480/runat_15102003_131913_20548/app_query.fasta_l.4685
-DB=est -OFMT=fastap -SUFFIX=txt -MINMATCH=0.1 -LOOPCPI=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DLOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480.@CGN_l.12645-runat_15102003_131913_20548 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRESH=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database : EST:★

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: qb\_gss1:

29: qb\_qss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	326	100.0	378	14	CB266492	CB266492 1005398 H
2	326	100.0	423	12	BM703962	BM703962 UI-E-CK1-
3	326	100.0	464	2	BX477782	BX477782 Homo sapi
4	326	100.0	531	12	BM830170	BM830170 K-EST0103
5	326	100.0	536	14	CB129403	CB129403 K-EST0179
6	326	100.0	547	12	BM787824	BM787824 K-EST0066
7	326	100.0	574	12	BM829991	BM829991 K-EST0103
8	326	100.0	691	14	CB853376	CB853376 UI-CF-FNO
9	326	100.0	849	14	CA489799	CA489799 AGENCOURT
10	326	100.0	852	14	CD516283	CD516283 AGENCOURT
11	326	100.0	885	14	CA455074	CA455074 AGENCOURT
12	326	100.0	893	14	CA455141	CA455141 AGENCOURT
13	326	100.0	894	14	CA454131	CA454131 AGENCOURT
14	326	100.0	899	14	CD558386	CD558386 AGENCOURT
15	326	100.0	916	13	BU845369	BU845369 AGENCOURT
16	326	100.0	919	13	BU845391	BU845391 AGENCOURT
17	326	100.0	932	14	CA487981	CA487981 AGENCOURT
18	326	100.0	944	14	CD558559	CD558559 AGENCOURT
19	326	100.0	1004	14	CD515356	CD515356 AGENCOURT
20	326	100.0	4715	11	AF1318349	AF1318349 Homo sapi
21	322	98.8	685	9	AU123871	AU123871 AU123871
22	313	96.0	924	14	CA489100	CA489100 AGENCOURT
23	313	96.0	984	14	CA489084	CA489084 AGENCOURT
24	309	94.8	866	14	CA488343	CA488343 AGENCOURT
25	302	92.6	348	9	A1909847	A1909847 QV-FET225-
26	300	92.0	444	12	BM789286	BM789286 K-EST0068
27	300	92.0	518	10	BF039952	BF039952 BP250015A
28	295	90.5	537	12	BM793744	BM793744 K-EST0074
29	292	89.6	505	12	BM854421	BM854421 K-EST0136
30	283	86.8	294	10	BF757857	BF757857 CM4 -CT057
31	283	86.8	674	10	BG675003	BG675003 602621225
32	279	85.6	527	14	CB719751	CB719751 AMGNUG:N
33	277	85.0	496	12	BM795191	BM795191 K-EST0076
34	276	84.7	422	10	BF554233	BF554233 UI-R-CO-h
35	276	84.7	904	13	BU931894	BU931894 AGENCOURT
36	275	84.4	630	2	BSM073298	BX4831330 Homo sapi
37	269	82.5	568	13	BQ831943	BQ831943 LL6in1122
38	269	82.5	577	13	BQ829915	BQ829915 LL6in12239
39	269	82.5	592	10	BE382065	BE382065 601272225
40	269	82.5	629	10	BE374596	BE374596 601272225
41	269	82.5	715	13	BU052728	BU052728 UI-M-EX0-
42	269	82.5	726	12	B1648082	B1648082 603279290
43	269	82.5	770	12	B1645946	B1645946 603274921
44	269	82.5	773	12	B1157032	B1157032 602921306
45	269	82.5	785	12	B1688650	B1688650 603311262

## ALIGNMENTS

RESULT 1	CB266492	378 bp	mrna	linear	EST 20-FEB-2003
LOCUS	CB266492	1005398	Human Fat Cell 5'-Stretch Plus cDNA Library	Homo sapiens	
DEFINITION	CB266492	cDNA 5', mRNA sequence.			
ACCESSION	CB266492				
VERSION	CB266492.1	GI:28441078			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 378)				

# **AUTHORS** **TITLE** **JOURNAL** **COMMENT**

Yang, R.-Z., Shuldiner, A. and Gong, D.-W.  
EST analysis of human adipose gene expression  
Unpublished  
Contact: Gong Da-Wei  
Division of Endocrinology, Diabetes and Nutrition  
University of Maryland  
660 Redwood St, HH497, Baltimore, MD 21201, USA  
Tel: 410 706 1672  
Fax: 410 706 1622  
Email: dgong@medicine.umaryland.edu  
PCR primers  
FORWARD: CTCGGGAAGCGCCGATGCTGTGGT  
BACKWARD: AATAGACTACTATAGGCGGAATTGG  
Seq primer: GTTGGTACCGGAATTC.

## **FEATURES**

source

1..378

Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/sex="Male and Female"  
/tissue\_type="Adipose"  
/clone\_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"  
/note="Vector: lambda triplex"

BASE COUNT 71 a 110 c 117 g 80 t

## **Alignment Scores:**

Pred. No.: 1.72e-29 Length: 378  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

SEQ5 (1-59) x CB266492 (1-378)

QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
Db 133 CAGAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 192  
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40  
Db 193 GAGGACGATGACATGGGGACCTGTGGATGCTGAGGAGTATCTGTGTACCCAGCAGGCG 252  
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59  
Db 253 TTCCTCTGTCCAGACCTGCGCGCGCTGGGGGCGATGTCACACAGGACCCGC 309

## **RESULT 2**

BM703962  
LOCUS 423 bp mRNA linear EST 28-FEB-2002  
DEFINITION UI-E-CK1-afk-o-11-0-UI-ri UI-E-CK1 Homo sapiens cDNA clone  
UI-E-CK1-afk-o-11-0-UI 5', mRNA sequence.

ACCESSION BM703962

VERSION 1

KEYWORDS Normalization and subtraction: two approaches to facilitate gene

SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Bonaldo, M.F., Lennon, G. and Soares, M.B.

discovery

Genome Res. 6 (9), 791-806 (1996)

9704447

889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEERF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 Reverse.

## **FEATURES**

source

1..423

Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="UI-E-CK1-afk-o-11-0-UI"  
/tissue\_type="Retina Foveal and Macular"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-CK1"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-E-CK1 is a normalized cDNA library containing the  
following tissue(s): Retina Foveal and Macular. The  
library was constructed according to Bonaldo, Lennon and  
Soares, Genome Research, 6:791-806, 1996. First strand  
containing a Not I site. Double stranded cDNA was ligated  
to an EcoR I adaptor, digested with Not I, and cloned  
directionally into pT7T3-Pac vector. The oligonucleotide  
used to prime the synthesis of first-strand cDNA contains  
a library tag sequence that is located between the Not I  
site and the (dT)18 tail. The sequence tag for this  
library is GTCC. This library was created for the program,  
Gene Discovery in the Visual System, supported by National  
Eye Institute (NEI)." 85 t

BASE COUNT 78 a 129 c 131 g

## **ORIGIN**

Alignment Scores:  
Pred. No.: 1.95e-29 Length: 423  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

SEQ5 (1-59) x BM703962 (1-423)

QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
Db 156 CAGAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 215  
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40  
Db 216 GAGGACGATGACATGGGGACCTGTGGATGCTGAGGAGTATCTGTGTACCCAGCAGGCG 275  
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59  
Db 276 TTCCTCTGTCCAGACCTGCGCGCGCTGGGGGCGATGTCACACAGGACCCGC 332

## **RESULT 3**

HSN067221

ID HSN067221 standard; RNA; EST; 464 BP.

XX BX477782;

AC BX477782.1

XX BX477782.1

SV BX477782.1

XX BX477782.1

DT 09-MAY-2003 (Rel. 75, Created)

DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)

XX Homo sapiens mRNA; EST DKFZp686M12198\_r1 (from clone DKFZp686M12198)

DE Homo sapiens mRNA; EST DKFZp686M12198\_r1 (from clone DKFZp686M12198)

XX Homo sapiens mRNA; EST DKFZp686M12198\_r1 (from clone DKFZp686M12198)

KW EST; expressed sequence tag.

XX Homo sapiens (human)  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;  
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
XX [1]  
RN 1-464  
RP Bloecker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,  
RA Han M., Wiemann S.;  
RT Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.  
RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY  
XX  
CC This is the 5' sequence of the clone insert  
CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
CC Sequenced by GBF (National Research Centre for Biotechnology  
CC Ltd., Braunschweig/Germany) within the cDNA sequencing  
CC consortium of the German Genome Project.  
CC No sl sequence available.  
CC This clone (DKFZp686M12198) is available at the RZPD in Berlin.  
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,  
CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de  
XX  
FH Key Location/Qualifiers  
FH 1. 464  
FH /db\_xref="taxon:9606"  
FH /mol\_type="mRNA"  
FH /organism="Homo sapiens"  
FH /clone="DKFZp686M12198"  
FH /clone.lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host  
FH DH10B; sites SfiIA + SfiIB"  
FH /dev\_stage="adult"  
FH /tissue\_type="CDNA-collection"  
XX  
SQ Sequence 464 BP; 87 A; 127 C; 149 G; 101 T; 0 other;

Alignment Scores:  
Pred. No.: 2,15e-29 Length: 464  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

SEQ5 (1-59) x HSM067221 (1-464)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
|||||  
Db 257 CAGATGAGGACTTGGGCCAGCCAGTCCCTGGGACAGCACCCTTACCGCTCACTGCTG 316

QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
|||||  
Db 317 GAGGACGATGACATGGGGACCTGGTGGTGTGAGGAGTATCTGGTACCCAGCAGGCG 376

QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59  
|||||  
Db 377 TTCCTTCTGCCAGACCTTGGCCCGGCGCTGGGGCGATGGTCCACACGACGCCG 433

RESULT 4  
BM830170  
LOCUS BM830170 531 bp mRNA linear EST 06-MAR-2002  
DEFINITION K-EST0103592 S21SNU520s1 Homo sapiens cDNA clone S21SNU520s1-19-D03  
5', mRNA sequence.  
ACCESSION BM830170  
VERSION BM830170.1 GI:19186579  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 531)

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 19 row: D column: 03  
High quality sequence stop: 531.  
Location/Qualifiers  
1. 531  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S21SNU520s1-19-D03"  
/sex="F"  
/tissue\_type="Stomach"  
/cell\_type="Floating aggregates"  
/lab\_host="SNU-520"  
/cell\_line="Top10F/"  
/clone.lib="S21SNU520s1"  
/notes="Organ: Stomach; Vector: pTZ18RP1; Site\_1: EcoRI;  
Site\_2: NotI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including EcoR  
I site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized from oligo dT-selected mRNA by  
priming with dT-tailed vector. The dT-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F' by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library. After analyzing and  
sequencing about 2,000 ~ 3,000 colonies in original cDNA  
library, the abundant cDNAs were selected and amplified by  
PCR reaction using vector region primer including T7  
promotor as 5' primer and N(dT)14 as 3' primer. The PCR  
products were used as template for synthesis of  
biotinylated single stranded RNA by in vitro transcription  
reaction. The synthesized RNA probes were hybridized with  
antisense single stranded cDNAs prepared from original  
library and incubated with avidin-gel. After removing  
DNA-RNA hybrids by centrifuge, the subtracted cDNA  
libraries were constructed by transformation of the  
remaining DNA into competent cells E. coli Top10F' with  
electroporation method."

BASE COUNT 102 a 149 c 166 g 114 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,5e-29 Length: 531  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

SEQ5 (1-59) x BM830170 (1-531).

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
|||||  
Db 343 CAGATGAGGACTTGGGCCAGCCAGTCCCTGGGACAGCACCCTTACCGCTCACTGCTG 402

QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40

```

|||||
403 GAGGACCATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGC 462
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
463 TTCTTCTGTCCAGACCTGCCCGGGCGTGGGGGCGATGTCACACAGGCCCGC 519

RESULT 5
LOCUS CB129403 536 bp mRNA linear EST 29-JAN-2003
DEFINITION K-EST0179037 C1SNUI7 Homo sapiens cDNA clone C1SNUI7-5-C08 5', mRNA
sequence.
ACCESSION CB129403
VERSION CB129403.1 GI:28093000
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 536)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yusong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 5 row: C column: 08
High quality sequence stop: 536.
Location/Qualifiers
1..536
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="C1SNUI7-5-C08"
/sex="F"
/tissue_type="Uterine"
/cell_type="Epithelial"
/lab_host="SNU-17"
/clone_lib="C1SNUI7"
/notes="Organ: Cervix; Vector: pCNS-D2; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including
EcoRI site by treatment of T4 RNA ligase and the first
strand cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10r' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
BASE COUNT 105 a 150 c 167 g 114 t
ORIGIN

Alignment Scores:
Pred. No.: 2,52e-29 Length: 536
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

```

```

SEQ5 (1-59) x CB129403 (1-536)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
348 CAGAAATGAGACTTGGGCCAGCCAGTCCCTTGACAGCACCCTTACCCCTCACTGCTG 407
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
408 GAGGACGATCACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGC 467
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
468 TTCTTCTGTCCAGACCTGCCCGGGCGTGGGGGCGATGTCACACAGGCCCGC 524

RESULT 6
LOCUS BM787824 547 bp mRNA linear EST 05-MAR-2002
DEFINITION K-EST0066898 S1LSNU1 Homo sapiens cDNA clone S1LSNU1-23-G01 5',
mRNA sequence.
ACCESSION BM787824
VERSION BM787824.1 GI:19136056
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yusong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 23 row: G column: 01
High quality sequence stop: 547.
Location/Qualifiers
1..547
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1LSNU1-23-G01"
/sex="M"
/tissue_type="Stomach"
/cell_type="Lymphoblast-like"
/lab_host="SNU-1"
/clone_lib="S1LSNU1"
/notes="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI;
Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including Sfil
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using Sfil
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with Sfil and
cloned into DraIII- digested pME18S-FL3 vector. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10r' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
BASE COUNT 102 a 187 c 163 g 95 t
ORIGIN

Alignment Scores:
Pred. No.: 2,58e-29 Length: 547

```



Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

SEQ5 (1-59) x BM87824 (1-547)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 |||||  
 Db 48 CAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGACCTTCTACCGCTCACTGCTG 107  
 |||||  
 Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
 |||||  
 Db 108 GAGGACGATGACATGGGGGACCTGGTGATCTGAGGAGTATCTGGTACCCACGACGGC 167  
 |||||  
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59  
 |||||  
 Db 168 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCGATGTTCCACACGACGACCGC 224  
 |||||

RESULT 7  
 BM829991  
 LOCUS  
 DEFINITION K-EST0103390 S21SNU520s1 Homo sapiens cDNA clone S21SNU520s1-17-C05  
 5', mRNA sequence.  
 ACCESSION BM829991  
 VERSION BM829991.1 GI:19186400  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 574)  
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@kribb.re.kr  
 Plate: 17 row: C column: 05  
 High quality sequence stop: 574.  
 Location/Qualifiers  
 1..574  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S21SNU520s1-17-C05"  
 /sex="F"  
 /tissue\_type="Stomach"  
 /cell\_type="Floating aggregates"  
 /cell\_line="SNU-520"  
 /lab\_host="Top10F"  
 /clone\_lib="S21SNU520s1"  
 /note="Organ: Stomach; Vector: pTZ18Rpl; Site:1: EcoRI;  
 Site:2: NotI; The poly (A)+ RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then decapped  
 with tabacco acid pyrophosphatase (TAP). The decapped  
 intact mRNA was ligated with DNA-RNA linker including EcoR  
 I site by treatment of T4 RNA ligase and the first strand  
 cDNA was synthesized from oligo dt-selected mRNA by  
 priming with dt-tailed vector. The dt-tailed vector was  
 adjusted to have about 60nt. The cDNA vector was  
 circularized with E. coli DNA ligase after digestion of  
 EcoRI which site is also included in vector. An RNA strand  
 converted to a DNA strand by Okayama-Berg method. The  
 obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10F' by electroporation method.

21C Frontier Korean EST Project 2001

Unpublished

COMMENT

Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@kribb.re.kr  
 Plate: 17 row: C column: 05  
 High quality sequence stop: 574.

FEATURES

source

The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(dt)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F' with electroporation method.

BASE COUNT 109 a 159 c 185 g 121 t  
 ORIGIN

Alignment Scores:

Pred. No.: 2.72e-29 Length: 574  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

SEQ5 (1-59) x BM829991 (1-574)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 |||||  
 Db 343 CAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGACCTTCTACCGCTCACTGCTG 402  
 |||||  
 Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
 |||||  
 Db 403 GAGGACGATGACATGGGGGACCTGGTGATCTGAGGAGTATCTGGTACCCACGACGGC 462  
 |||||  
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59  
 |||||  
 Db 463 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCGATGTTCCACACGACGACCGC 519  
 |||||

RESULT 8

CB853376/c

LOCUS

DEFINITION UI-CF-FN0-agd-h-18-0-UI.s1 UI-CF-FN0 Homo sapiens cDNA clone  
 UI-CF-FN0-agd-h-18-0-UI 3', mRNA sequence.

ACCESSION CB853376

VERSION CB853376.1 GI:30048540

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 691)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: McCray, PB

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com) or from Open Biosystems

(www.openbiosystems.com).

Seq primer: M13 FORWARD  
POLYA=No.

## FEATURES

source

Location/Qualifiers

```
1. .691
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="UI-CF-FNO-aggd-h-18-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="PH10B (Life Technologies) (T1 phase resistant)"
/clone_lib="UI-CF-FNO"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: Ecor I; Site_2: Not I; UI-CF-FNO is a subtracted cDNA library derived from two normalized Human lung epithelial cell libraries (EN1 and DUL) The library was subtracted according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bent-soares@uiowa.edu
TAG_SEQ=None found"
```

BASE COUNT 140 a 224 c 191 g 131 t 5 others  
ORIGIN

Alignment Scores:  
Pred. No.: 3,32e-29 Length: 691  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

SEQ5 (1-59) x CB853376 (1-691)

```
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 317 CAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCG 258
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
|||||
Db 257 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCAGCAGGCG 198
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
|||||
Db 197 TTCTTCTGTCAGACCTGCCCGGGCGCTGGGGCATGCTGCACACAGGACCGC 141
```

RESULT 9  
CA489799  
LOCUS  
DEFINITION AGENCOURT\_10810428 MAPcL Homo sapiens cDNA clone IMAGE:6722285 5',  
mRNA sequence.  
ACCESSION CA489799  
VERSION  
KEYWORDS  
SOURCE EST. GI:24952590  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 849)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Kristi A. Eglund, Ira Pastan  
CDNA Library Preparation: Invitrogen Corp  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM14284 row: m column: 05  
High quality sequence stop: 539.  
Location/Qualifiers

## FEATURES

source

```
1. .849
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:6722285"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1", LNCap"
/lab_host="EMDH10B"
/clone_lib="MAPcL"
/note="Vector: pCMV-SPORT6; Site_1: EcorV; Site_2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan. Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."
```

BASE COUNT 179 a 272 c 243 g 155 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4,16e-29 Length: 849  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

SEQ5 (1-59) x CA489799 (1-849)

```
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 185 CAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCG 244
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
|||||
Db 245 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCAGCAGGCG 304
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
|||||
Db 305 TTCTTCTGTCAGACCTGCCCGGGCGCTGGGGCATGCTGCACACAGGACCGC 361
```

RESULT 10  
CD516283

LOCUS  
DEFINITION AGENCOURT\_14364504 NIH\_MGC\_181 Homo sapiens cDNA clone  
IMAGE:30408904 5', mRNA sequence.  
ACCESSION CD516283  
VERSION  
KEYWORDS  
SOURCE EST. GI:31448001  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 852)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Dr. Michael Brownstein  
CDNA Library Preparation: Invitrogen Corp  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: NDAM492 row: j column: 17  
High quality sequence stop: 662.  
Location/Qualifiers

## FEATURES

## source

```

1..852
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30408904"
/tissue_type="White Matter"
/dev_stage="Unknown"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_181"
/Note="Vector: PCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
(destroyed); Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(invitrogen). Note: this is a NIH_MGC Library."
BASE COUNT 159 a 288 c 247 g 156 t 2 others
ORIGIN

```

## Alignment Scores:

```

Pred. No.: 4.17e-29 Length: 852
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

```

SEQ5 (1-59) x CD516283 (1-852)

```

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 95 CAGAATGAGGACTTGGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 154
|||||
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
|||||
Db 155 GAGGACGATGACATGGGGGAGCTGGTGGATGCTGAGGAGATCTGTTACCCAGCAGGCGC 214
|||||
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
Db 215 TTCTTCTGTCAGACCTGCGCCGGCGCTGGGGCGATGGTCCACACGAGCAGCGC 271
|||||

```

## RESULT 11

```

CA455074
LOCUS CA455074 885 bp mRNA linear EST 12-NOV-2002
DEFINITION AGENCOURT_10735919 MAPcL Homo sapiens cDNA clone IMAGE:6722585 5',
mRNA sequence.
ACCESSION CA455074
VERSION CA455074.1 GI:24905427
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL 1 (bases 1 to 885)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov

```

```

CNA Library Preparation: Invitrogen Corp
CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14285 row: i column: 17
High quality sequence stop: 717.

```

## FEATURES

## source

```

1..885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722585"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1"

```

```

' LNCap"
/lab_host="EMDH10B"
/clone_lib="MAPcL"
/Note="Vector: PCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
Subtracted with brain, liver, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bunkook Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."
BASE COUNT 178 a 297 c 261 g 149 t
ORIGIN

```

## Alignment Scores:

```

Pred. No.: 4.35e-29 Length: 885
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

```

SEQ5 (1-59) x CA455074 (1-885)

```

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 1 CAGAATGAGGACTTGGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 60
|||||
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
|||||
Db 61 GAGGACGATGACATGGGGGAGCTGGTGGATGCTGAGGAGATCTGTTACCCAGCAGGCGC 120
|||||
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
Db 121 TTCTTCTGTCAGACCTGCGCCGGCGCTGGGGCGATGGTCCACACGAGCAGCGC 177
|||||

```

## RESULT 12

```

CA455141
LOCUS CA455141 893 bp mRNA linear EST 12-NOV-2002
DEFINITION AGENCOURT_10735980 MAPcL Homo sapiens cDNA clone IMAGE:6722663 5',
mRNA sequence.
ACCESSION CA455141
VERSION CA455141.1 GI:24905561
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
1 (bases 1 to 893)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov

```

```

Tissue Procurement: Kristi A. Eglund, Ira Pastan
CNA Library Preparation: Invitrogen Corp
CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14285 row: 1 column: 23
High quality sequence stop: 682.

```

## FEATURES

## source

```

1..893
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722663"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1"
' LNCap"
/clone_lib="MAPcL"

```

/note="Vector: pCMV-SPORT6; Site\_1: EcoRV; Site\_2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

BASE COUNT 175 a 302 c 261 g 155 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4,39e-29 Length: 893  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

SEQ5 (1-59) x CA455141 (1-893)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
|||||  
Db 67 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTCTG 126  
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40  
|||||  
Db 127 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCAGCAGGCG 186  
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59  
|||||  
Db 187 TTCTTCTGTCAGACCTGCCCCGGCGCTGGGGCATGCTCCACACAGGACCCGC 243

RESULT 13

CA454131

LOCUS CD454131 894 bp mRNA linear EST 12-NOV-2002  
DEFINITION AGENCOURT\_10738550 MAPcL Homo sapiens cDNA clone IMAGE:6718792 5', mRNA sequence.

ACCESSION CA454131

VERSION CA454131.1 GI:24903569

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 894)

NIH-MGC http://mgi.nci.nih.gov/

Unpublished

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Kristi A. Eglund, Ira Pastan

CDNA Library Preparation: Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM14275 row: k column: 16

High quality sequence stop: 636.

Location/Qualifiers

1. .894

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6718792"

/cell\_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1

, INCap"

/lab\_host="EMDH10B"

/clone\_lib="MAPcL"

/note="Vector: pCMV-SPORT6; Site\_1: EcoRV; Site\_2: Not I;

Subtracted with brain, liver, lung, kidney and muscle.

Directionally cloned. Priming method: oligo-dT. Average

insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

BASE COUNT 180 a 303 c 260 g 151 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4,4e-29 Length: 894  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

SEQ5 (1-59) x CA454131 (1-894)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
|||||  
Db 24 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTCTG 83  
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40  
|||||  
Db 84 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCAGCAGGCG 143  
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59  
|||||  
Db 144 TTCTTCTGTCAGACCTGCCCCGGCGCTGGGGCATGCTCCACACAGGACCCGC 200

RESULT 14

CD558386

LOCUS CD558386

DEFINITION AGENCOURT\_14476827 NIH\_MGC\_181 Homo sapiens cDNA clone

IMAGE:30396412 5', mRNA sequence.

ACCESSION CD558386

VERSION CD558386.1 GI:31584454

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 899)

NIH-MGC http://mgi.nci.nih.gov/

Unpublished

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein

CDNA Library Preparation: Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: NDAM478 row: b column: 05

High quality sequence start: 47

High quality sequence stop: 443.

Location/Qualifiers

1. .899

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:30396412"

/tissue\_type="White Matter"

/dev\_stage="Unknown"

/lab\_host="DH10B-Ton A ( T1 and T5 phage resistances)"

/clone\_lib="NIH\_MGC\_181"

/note="Vector: pCMV-SPORT6.1; Site\_1: NotI; Site\_2: EcoRV

(destroyed); Library is oligo-dT primed and directionally

cloned (EcoRV site is destroyed upon cloning). Average insert size 1.42 kb. Library was constructed by (Invitrogen). Note: this is a NIH\_MGC Library."

BASE COUNT 165 a 338 c 237 g 159 t

## ORIGIN

## Alignment Scores:

Pred. No.: 4.42e-29 Length: 899  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

SEQ5 (1-59) x CD558386 (1-899)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
|||||  
Db 74 CAGAAATGAGGACTTGGGCCCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCACTGCTG 133  
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
|||||  
Db 134 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCGC 193  
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59  
|||||  
Db 194 TTCTTCTGTCCAGACCTTGCCTGGGGCGCTGGGGCGATGGTCCACACAGCACCGC 250

## RESULT 15

BU845369

LOCUS

DEFINITION BU845369 916 bp mRNA linear EST 16-OCT-2002  
IMAGE:6578980 5', mRNA sequence.

ACCESSION BU845369

VERSION BU845369.1 GI:24029810

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 916)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1CM2781 row: j column: 04

High quality sequence stop: 654.

## FEATURES

source

1, 916  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6578980"  
/tissue\_type="teratocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_109"

/note="Organ: ovary; Vector: pOTB7; Site\_1: EcoRI; Site\_2:  
XhoI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GCCACGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

BASE COUNT 177 a 288 c 274 g 177 t

## ORIGIN

Alignment Scores:  
Pred. No.: 4.51e-29 Length: 916  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

SEQ5 (1-59) x BU845369 (1-916)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
|||||  
Db 343 CAGAAATGAGGACTTGGGCCCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCACTGCTG 402  
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
|||||  
Db 403 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCGC 462  
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59  
|||||  
Db 463 TTCTTCTGTCCAGACCTTGCCTGGGGCGCTGGGGCGATGGTCCACACAGCACCGC 519

Search completed: October 16, 2003, 17:04:51

Job time : 463.761 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 16, 2003, 11:08:07 : Search time 60.9885 Seconds  
(without alignments)  
2540.503 Million cell updates/sec

Title: SEQ5

Perfect score: 326

Sequence: 1 QNEDLGASPDLSTFYRSL.....GFFCPDPAPGAGGVHRRH 59

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool\_p/HOLLERAN80/runat\_15102003\_131915\_20662/app\_query.fasta\_1.4685  
-DB=PublishedApplications\_NA -OFFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USG=HOLLERAN80 -CGN 1.1.397 -runat\_15102003\_131915\_20662  
-NCPU=6 -ICPU=1 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	326	100.0	1115	14	US-10-102-806-165 Sequence 165, App

2	326	100.0	1713	12	US-10-378-393-14	Sequence 14, Appl
3	326	100.0	1755	10	US-09-930-125-6	Sequence 6, Appl
4	326	100.0	1767	10	US-09-930-125-4	Sequence 4, Appl
5	326	100.0	1773	10	US-09-930-125-7	Sequence 7, Appl
6	326	100.0	1806	10	US-09-930-125-5	Sequence 5, Appl
7	326	100.0	2411	12	US-10-378-393-10	Sequence 10, Appl
8	326	100.0	3765	12	US-10-207-498-5	Sequence 5, Appl
9	326	100.0	3768	9	US-09-811-123-8	Sequence 8, Appl
10	326	100.0	3768	9	US-09-811-115-2	Sequence 2, Appl
11	326	100.0	3768	10	US-09-854-356-9	Sequence 9, Appl
12	326	100.0	3768	10	US-09-930-125-1	Sequence 1, Appl
13	326	100.0	3768	12	US-10-313-644-1	Sequence 1, Appl
14	326	100.0	4473	11	US-09-441-411-5	Sequence 5, Appl
15	326	100.0	4473	12	US-10-101-510-81	Sequence 81, Appl
16	326	100.0	4473	14	US-10-146-473-32	Sequence 32, Appl
17	326	100.0	4473	14	US-10-207-655-44	Sequence 44, Appl
18	326	100.0	4530	10	US-09-877-177-11	Sequence 11, Appl
19	326	100.0	4530	12	US-10-007-926A-119	Sequence 119, App
20	326	100.0	4530	12	US-10-101-510-124	Sequence 124, App
21	326	100.0	4530	12	US-10-338-730-1	Sequence 1, Appl
22	326	100.0	4530	14	US-10-177-293-125	Sequence 125, App
23	326	100.0	4543	10	US-09-769-508-1	Sequence 1, Appl
24	326	100.0	4606	12	US-09-971-392-70	Sequence 70, Appl
25	326	100.0	4642	14	US-10-198-846-10896	Sequence 10896, A
26	326	100.0	9274	9	US-09-811-123-7	Sequence 7, Appl
27	326	100.0	9274	9	US-09-811-115-1	Sequence 1, Appl
28	279	85.6	3955	10	US-09-870-759-117	Sequence 117, App
29	279	85.6	3955	10	US-09-854-356-10	Sequence 10, Appl
30	279	85.6	3955	12	US-09-751-708A-117	Sequence 117, App
31	269	82.5	3771	10	US-09-854-356-11	Sequence 11, Appl
32	154	47.2	249	10	US-09-960-352-4186	Sequence 4186, Ap
33	138	42.3	3633	9	US-09-725-433-1	Sequence 1, Appl
34	138	42.3	5264	10	US-09-920-300A-1731	Sequence 1731, Ap
35	138	42.3	5264	12	US-10-039-926-1731	Sequence 1731, Ap
36	138	42.3	5264	13	US-10-033-528-1731	Sequence 137, Ap
37	138	42.3	5532	12	US-10-007-926A-137	Sequence 95, Appl
38	138	42.3	5532	12	US-10-101-510-95	Sequence 95, Appl
39	138	42.3	10058	10	US-09-744-298-98	Sequence 98, Appl
40	138	42.3	197496	10	US-09-877-177-10	Sequence 10, Appl
41	130	39.9	1935	12	US-10-259-451-1	Sequence 1, Appl
42	108.5	33.3	2437	11	US-09-158-722-3	Sequence 3, Appl
43	108.5	33.3	5484	10	US-09-940-101-1	Sequence 1, Appl
44	108.5	33.3	5484	14	US-10-207-655-46	Sequence 46, Appl
45	98	30.1	450	12	US-10-007-926A-136	Sequence 136, App

ALIGNMENTS

RESULT 1

US-10-102-806-165  
; Sequence 165, Application US/10102806  
; Publication No. US20030054421A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA103p1c1  
; CURRENT APPLICATION NUMBER: US/10/102,806  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 846  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 165  
; LENGTH: 1115  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (390)

```
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-102-806-165

Alignment Scores:
Pred. No.: 1.23e-38 Length: 1115
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

SEQ5 (1-59) x US-10-102-806-165 (1-1115)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 455 CAGAATGAGGACTGGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTCTG 514
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 515 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCACGAGGCG 574
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 575 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCGATGCTGCCACACAGGACCGC 631

RESULT 2
US-10-378-393-14
; Sequence 14, Application US/10378393
; Publication NO. US20030182668A1
; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVATED
; FILE REFERENCE: D0254 NP
; CURRENT APPLICATION NUMBER: US/10/378,393
; PRIOR FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,889
; PRIORITY FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-378-393-14

Alignment Scores:
Pred. No.: 2.05e-38 Length: 1713
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ5 (1-59) x US-10-378-393-14 (1-1713)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 913 CAGAATGAGGACTGGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTCTG 972
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 973 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCACGAGGCG 1032
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 1032 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCGATGCTGCCACACAGGACCGC 1122

; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-930-125-6

Alignment Scores:
Pred. No.: 2.11e-38 Length: 1755
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ5 (1-59) x US-09-930-125-6 (1-1755)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 946 CAGAATGAGGACTGGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTCTG 1005
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 1006 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCACGAGGCG 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 1066 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCGATGCTGCCACACAGGACCGC 1122

RESULT 3
US-09-930-125-6
; Sequence 6, Application US/09930125
; Publication NO. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-6

Alignment Scores:
Pred. No.: 2.11e-38 Length: 1755
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ5 (1-59) x US-09-930-125-6 (1-1755)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 946 CAGAATGAGGACTGGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTCTG 1005
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 1006 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCACGAGGCG 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 1066 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCGATGCTGCCACACAGGACCGC 1122

RESULT 4
US-09-930-125-4
; Sequence 4, Application US/09930125
; Publication NO. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-4
```



## Alignment Scores:

Pred. No.: 2,13e-38 Length: 1767  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps:

SEQ5 (1-59) x US-09-930-125-4 (1-1767)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 |||||  
 Db 946 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1005  
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
 |||||  
 Db 1006 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 1065  
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArg 59  
 |||||  
 Db 1066 TTCTTCTGTCCAGACCTGCCCGGCGCTGGGGCGATGGTCCACCAGCAGCACCGC 1122

## RESULT 5

US-09-930-125-7

; Sequence 7, Application US/09930125

; Publication No. US20020193329A1

## GENERAL INFORMATION:

; APPLICANT: Hand-Zimmerman, Susan  
 ; APPLICANT: Cheever, Martin A.  
 ; APPLICANT: Foy, Teresa M.  
 ; APPLICANT: Lodes, Michael J.  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: McNeill, Patricia D.  
 ; APPLICANT: Vedvick, Thomas S.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS

; FILE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES

; CURRENT APPLICATION NUMBER: US/09/930,125

; CURRENT FILING DATE: 2001-08-14

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 7

; LENGTH: 1773

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-930-125-7

## Alignment Scores:

Pred. No.: 2,14e-38 Length: 1773  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps:

SEQ5 (1-59) x US-09-930-125-7 (1-1773)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 |||||  
 Db 970 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1029  
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
 |||||  
 Db 1030 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 1089  
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArg 59  
 |||||  
 Db 1090 TTCTTCTGTCCAGACCTGCCCGGCGCTGGGGCGATGGTCCACCAGCAGCACCGC 1146

## RESULT 6

US-09-930-125-5

; Sequence 5, Application US/09930125

; Publication No. US20020193329A1

## GENERAL INFORMATION:

; APPLICANT: Hand-Zimmerman, Susan  
 ; APPLICANT: Cheever, Martin A.  
 ; APPLICANT: Foy, Teresa M.  
 ; APPLICANT: Lodes, Michael J.  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: McNeill, Patricia D.  
 ; APPLICANT: Vedvick, Thomas S.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS

; FILE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES

; CURRENT APPLICATION NUMBER: US/09/930,125

; CURRENT FILING DATE: 2001-08-14

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 5

; LENGTH: 1806

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-930-125-5

## Alignment Scores:

Pred. No.: 2,19e-38 Length: 1806  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps:

SEQ5 (1-59) x US-09-930-125-5 (1-1806)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 |||||  
 Db 1003 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1062  
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
 |||||  
 Db 1063 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 1122  
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArg 59  
 |||||  
 Db 1123 TTCTTCTGTCCAGACCTGCCCGGCGCTGGGGCGATGGTCCACCAGCAGCACCGC 1179

## RESULT 7

US-10-378-393-10

; Sequence 10, Application US/10378393

; Publication No. US20030182668A1

## GENERAL INFORMATION:

; APPLICANT: Bol, David K.  
 ; APPLICANT: Carboni, Joan M.  
 ; APPLICANT: Rowley, Ronald B.  
 ; APPLICANT: Wong, Tai W.  
 ; APPLICANT: Lee, Francis Y.

; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVAT

; FILE OF INVENTION: TYROSINE KINASE RECEPTORS

; FILE REFERENCE: D0254 NP

; CURRENT APPLICATION NUMBER: US/10/378,393

; CURRENT FILING DATE: 2003-03-03

; PRIOR APPLICATION NUMBER: US 60/360,889

; PRIOR FILING DATE: 2002-03-01

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 10

; LENGTH: 2411

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-378-393-10

## Alignment Scores:

Pred. No.: 3,09e-38 Length: 2411  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

SEQ5 (1-59) x US-10-378-393-10 (1-2411)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 Db 1611 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1670  
 Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40  
 Db 1671 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCAGCAGCGC 1730  
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59  
 Db 1731 TTCTTCTGTCAGACCTGCCCGGGGCGCTGGGGGCGATGTTCCACACAGCACCGC 1787

#### RESULT 8

US-10-207-498-5  
 ; Sequence 5, Application US/10207498  
 ; Publication No. US20030143568A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Elizabeth Singer  
 ; APPLICANT: Ralf Landgraf  
 ; APPLICANT: Dennis J. Slamon  
 ; APPLICANT: David Eisenberg  
 ; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING  
 ; TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HEREGULIN AND HER3  
 ; FILE REFERENCE: 30448-103-US-U1  
 ; CURRENT APPLICATION NUMBER: US/10/207,498  
 ; PRIOR FILING DATE: 2002-07-29  
 ; PRIOR APPLICATION NUMBER: 60/308,431  
 ; PRIOR FILING DATE: 2001-07-27  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5  
 ; LENGTH: 3765  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)...(3765)  
 ; US-10-207-498-5

Alignment Scores:  
 Pred. No.: 5,25e-38 Length: 3765  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

SEQ5 (1-59) x US-10-207-498-5 (1-3765)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027  
 Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40  
 Db 3028 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCAGCAGCGC 3087  
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59  
 Db 3088 TTCTTCTGTCAGACCTGCCCGGGGCGCTGGGGGCGATGTTCCACACAGCACCGC 3144

#### RESULT 9

US-09-811-123-8  
 ; Sequence 8, Application US/09811123  
 ; Patent No. US20020001587A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sharon Erickson

; APPLICANT: Ralph Schwall  
 ; APPLICANT: Mark Sliwowski  
 ; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ETBB  
 ; TITLE OF INVENTION: ANTIBODY-MAYTANSINOID CONJUGATES  
 ; FILE REFERENCE: GENENT.073A2

; CURRENT APPLICATION NUMBER: US/09/811,123  
 ; PRIOR FILING DATE: 2001-03-16  
 ; PRIOR APPLICATION NUMBER: 60/238,327  
 ; PRIOR FILING DATE: 2000-10-05  
 ; PRIOR APPLICATION NUMBER: 09/602,530  
 ; PRIOR FILING DATE: 2000-06-23  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 8  
 ; LENGTH: 3768  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-811-123-8

#### Alignment Scores:

Pred. No.: 5,26e-38 Length: 3768  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

SEQ5 (1-59) x US-09-811-123-8 (1-3768)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027  
 Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40  
 Db 3028 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCAGCAGCGC 3087  
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59  
 Db 3088 TTCTTCTGTCAGACCTGCCCGGGGCGCTGGGGGCGATGTTCCACACAGCACCGC 3144

#### RESULT 10

US-09-811-115-2  
 ; Sequence 2, Application US/09811115  
 ; Patent No. US20020035736A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Erickson, Sharon  
 ; APPLICANT: Schwall, Ralph  
 ; APPLICANT: King, Kathleen  
 ; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL  
 ; FILE REFERENCE: GENENT.034A  
 ; CURRENT APPLICATION NUMBER: US/09/811,115  
 ; CURRENT FILING DATE: 2001-03-16  
 ; PRIOR APPLICATION NUMBER: 60/189,844  
 ; PRIOR FILING DATE: 2000-03-16  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 3768  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-811-115-2

Alignment Scores:  
 Pred. No.: 5,26e-38 Length: 3768  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

SEQ5 (1-59) x US-09-811-115-2 (1-3768)

```

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGATGAGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3088 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACAGCAGCGC 3144

RESULT 11
US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

Alignment Scores:
Pred. No.: 5,26e-38 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ5 (1-59) x US-09-854-356-9 (1-3768)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGATGAGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3088 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACAGCAGCGC 3144

RESULT 12
US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US2002019329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3765)
US-09-930-125-1

Alignment Scores:
Pred. No.: 5,26e-38 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ5 (1-59) x US-09-930-125-1 (1-3768)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGATGAGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3088 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACAGCAGCGC 3144

RESULT 13
US-10-313-644-1
; Sequence 1, Application US/10313644
; Publication No. US20030157119A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: Hand-Zimmerman, Susan
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
; FILE REFERENCE: 210121.483C3
; CURRENT APPLICATION NUMBER: US/10/313,644
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:

```

```

Db 3088 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACAGCAGCGC 3144

RESULT 12
US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US2002019329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3765)
US-09-930-125-1

Alignment Scores:
Pred. No.: 5,26e-38 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ5 (1-59) x US-09-930-125-1 (1-3768)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGATGAGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3088 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACAGCAGCGC 3144

RESULT 13
US-10-313-644-1
; Sequence 1, Application US/10313644
; Publication No. US20030157119A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: Hand-Zimmerman, Susan
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
; FILE REFERENCE: 210121.483C3
; CURRENT APPLICATION NUMBER: US/10/313,644
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:

```

```
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-313-644-1

Alignment Scores:
Pred. No.: 5,26e-38 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ5 (1-59) x US-10-313-644-1 (1-3768)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAATGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGACCTGTGGATGCTGAGGAGTATCTGTACCCAGCAGGCG 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 3088 TTCTTCTGTCCAGACCTGCCCCGGCGCTGGGGGCGATGCTCCACCACAGGACCGC 3144

RESULT 14
US-09-441-411-5
; Sequence 5, Application US/09441411
; Publication No. US20030008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441,411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-5

Alignment Scores:
Pred. No.: 6,45e-38 Length: 4473
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

SEQ5 (1-59) x US-09-441-411-5 (1-4473)
QY 1 GluAspGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGAAATGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3202 GAGGACGATGACATGGGGACCTGTGGATGCTGAGGAGTATCTGTACCCAGCAGGCG 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 3262 TTCTTCTGTCCAGACCTGCCCCGGCGCTGGGGGCGATGCTCCACCACAGGACCGC 3318

RESULT 15
US-10-101-510-81
; Sequence 81, Application US/10101510
; Publication No. US20030148295A1
```

```
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-81

Alignment Scores:
Pred. No.: 6,45e-38 Length: 4473
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ5 (1-59) x US-10-101-510-81 (1-4473)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGAAATGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3202 GAGGACGATGACATGGGGACCTGTGGATGCTGAGGAGTATCTGTACCCAGCAGGCG 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 3262 TTCTTCTGTCCAGACCTGCCCCGGCGCTGGGGGCGATGCTCCACCACAGGACCGC 3318

Search completed: October 17, 2003, 03:54:41
Job time : 65.9885 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 22:55:19 ; Search time 14.9563 Seconds  
(without alignments)  
1741.185 Million cell updates/sec

Title: SEQ5

Perfect score: 326

Sequence: 1 QNEDLGASPLDSTFYRSL.....GFFCPDPAPGAGGVHRRH 59

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delep 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O/cgn2\_1/USPTO\_spool\_P/HOLLERAN480/runat\_15102003\_131913\_20570/app\_query.fasta\_1.4685  
-DB-Issued\_Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 @cgn\_1\_1\_403 @runat\_15102003\_131913\_20570 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOG-100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	326	100.0	3768	2	US-08-625-101-1
2	326	100.0	3768	2	US-08-356-786-1
3	326	100.0	4473	2	US-09-048-804-1
4	326	100.0	4473	3	US-09-056-105-26
5	326	100.0	4530	1	US-08-229-515A-9
6	326	100.0	4530	1	US-08-645-865-9
7	326	100.0	4530	4	US-09-167-322-4
8	326	100.0	4530	4	US-09-527-487-1
9	326	100.0	4530	4	US-09-877-177A-11
10	279	85.6	3955	1	US-08-229-515A-14
11	279	85.6	3955	1	US-08-645-865-13
12	138	42.3	5532	2	US-08-475-035-3

13	138	42.3	5532	4	US-09-676-610B-17	Sequence 17, Appl
14	138	42.3	197496	4	US-09-877-177A-10	Sequence 10, Appl
15	136	41.7	410	3	US-08-604-991-11	Sequence 11, Appl
16	136	41.7	410	3	US-09-363-639-11	Sequence 11, Appl
17	136	41.7	576	1	US-08-783-275-3	Sequence 3, Appl
18	136	41.7	576	1	US-08-727-708-3	Sequence 3, Appl
19	136	41.7	576	2	US-08-766-677-1	Sequence 1, Appl
20	136	41.7	576	2	US-08-843-951-1	Sequence 1, Appl
21	108.5	33.3	2437	2	US-08-456-647B-3	Sequence 3, Appl
22	108.5	33.3	2437	2	US-08-237-401A-3	Sequence 3, Appl
23	108.5	33.3	3321	1	US-08-484-438-5	Sequence 5, Appl
24	108.5	33.3	5484	3	US-09-632-360A-3	Sequence 3, Appl
25	108.5	33.3	5501	1	US-08-484-438-1	Sequence 1, Appl
26	108.5	33.3	5555	1	US-08-484-438-3	Sequence 3, Appl
27	66	20.2	980	4	US-09-218-467B-6	Sequence 6, Appl
28	66	20.2	1209	4	US-09-170-496D-51	Sequence 51, Appl
29	66	20.2	1209	4	US-09-170-496D-191	Sequence 191, App
30	66	20.2	1269	3	US-09-224-426-1	Sequence 1, Appl
31	66	20.2	1269	3	US-09-478-601-1	Sequence 1, Appl
32	66	20.2	1269	3	US-09-478-602-1	Sequence 1, Appl
33	66	20.2	1316	3	US-08-602-809-1	Sequence 1, Appl
34	66	20.2	1316	5	PCT-US95-16472-1	Sequence 1, Appl
35	66	20.2	1385	3	US-08-984-288-1	Sequence 1, Appl
36	66	20.2	3488	4	US-09-218-467B-1	Sequence 1, Appl
37	65	19.9	1173	2	US-08-614-156B-4	Sequence 4, Appl
38	65	19.9	1173	3	US-09-213-864-1	Sequence 1, Appl
39	64	19.6	6462	4	US-09-792-594-10	Sequence 10, Appl
40	62.5	19.2	975	4	US-09-252-991A-12244	Sequence 12244, A
41	62.5	19.2	1518	4	US-09-252-991A-12104	Sequence 12104, A
42	62.5	19.2	2040	4	US-09-252-991A-12316	Sequence 12316, A
43	62	19.0	312	4	US-09-252-991A-16553	Sequence 16553, A
44	62	19.0	501	4	US-09-252-991A-10705	Sequence 10705, A
45	62	19.0	999	4	US-09-252-991A-15955	Sequence 15955, A

ALIGNMENTS

RESULT 1

US-08-625-101-1

; Sequence 1, Application US/08625101

; Patent No. 5863445

; GENERAL INFORMATION:

; APPLICANT: Cheever, Martin A.

; APPLICANT: Disis, Mary L.

; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE

; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION

; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu

; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/625,101

; FILING DATE: 01-APR-1996

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Sharkey, Richard G.

; REGISTRATION NUMBER: 32,629

; REFERENCE/DOCKET NUMBER: 920010.448C7

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 3768 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..3765  
US-08-625-101-1

Alignment Scores:  
Pred. No.: 3.24e-32 Length: 3768  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

SEQ5 (1-59) x US-08-625-101-1 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
DB 2968 CAGATGAGGAGCTTGGCCACGACCTCTTGGACAGCACCTTCTACCGCTCACTGTG 3027  
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyLeuValProGlnGlnGly 40  
DB 3028 GAGGACGATGACATGGGGACCTGTGGATGCTGAGGAGTATCTGTATCCCCAGCAGGCG 3087  
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59  
DB 3088 TTCTTCTGTCACAGACCTGCCCCGGCGCTGGGGCATGTGTCACACAGGACCGC 3144

## RESULT 2

US-08-356-786-1

; Sequence 1, Application US/08356786  
; Patent No. 5877305

## GENERAL INFORMATION:

; APPLICANT: Huston, James S.  
; APPLICANT: Oppermann, Hermann  
; APPLICANT: Houston, L. L.  
; APPLICANT: Ring, David B.  
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
; TITLE OF INVENTION: Marker  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
; STREET: Exchange Place, 53 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109

## COMPUTER READABLE FORM:

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/356,786  
; FILING DATE:

## CLASSIFICATION:

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/831,967  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pitcher, Edmund R.  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: CRP-053  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3768 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..3768  
OTHER INFORMATION: /note= "product = "cerB-b2"  
US-08-356-786-1

Alignment Scores:  
Pred. No.: 3.24e-32 Length: 3768  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

SEQ5 (1-59) x US-08-356-786-1 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
DB 2968 CAGATGAGGAGCTTGGCCACGACCTCTTGGACAGCACCTTCTACCGCTCACTGTG 3027  
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyLeuValProGlnGlnGly 40  
DB 3028 GAGGACGATGACATGGGGACCTGTGGATGCTGAGGAGTATCTGTATCCCCAGCAGGCG 3087  
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59  
DB 3088 TTCTTCTGTCACAGACCTGCCCCGGCGCTGGGGCATGTGTCACACAGGACCGC 3144

## RESULT 3

US-09-048-804-1

; Sequence 1, Application US/09048804  
; Patent No. 5968748

## GENERAL INFORMATION:

; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters  
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF  
; TITLE OF INVENTION: HUMAN HER-2 EXPRESSION  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748ris LLP  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: U.S.A.  
; ZIP: 19103

## COMPUTER READABLE FORM:

; MEDIUM TYPE: 1.44 Mb diskette  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/048,804  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Paul K. Legard  
; REGISTRATION NUMBER: 38,534  
; REFERENCE/DOCKET NUMBER: ISIS-2913  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4473 base pairs  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Unknown  
; ANTI-SENSE: NO

US-09-048-804-1

## Alignment Scores:

Pred. No.: 3,96e-32 Length: 4473  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

SEQ5 (1-59) x US-09-048-804-1 (1-4473)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
Db 3142 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTCTACCGCTCACTGCTG 3201  
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
Db 3202 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCACGACGAGGC 3261  
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59  
Db 3262 TTCTTCTGTCAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACGACGACCGC 3318

## RESULT 4

US-09-056-105-26

; Sequence 26, Application US/09056105

; Patent No. 6287569

; GENERAL INFORMATION:

; APPLICANT: KIPPS, THOMAS J.

; APPLICANT: WU, YUNQI

; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR

; FILE REFERENCE: 233/221

; CURRENT APPLICATION NUMBER: US/09/056,105

; EARLIER FILING DATE: 1998-04-06

; EARLIER APPLICATION NUMBER: 60/043,467

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 26

; LENGTH: 4473

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-056-105-26

## Alignment Scores:

Pred. No.: 3,96e-32 Length: 4473  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

SEQ5 (1-59) x US-09-056-105-26 (1-4473)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
Db 3142 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTCTACCGCTCACTGCTG 3201  
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
Db 3202 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCACGACGAGGC 3261  
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59  
Db 3262 TTCTTCTGTCAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACGACGACCGC 3318

## RESULT 5

US-08-229-515A-9

; Sequence 9, Application US/08229515A

; Patent No. 5518885

; GENERAL INFORMATION:

; APPLICANT: RAZIUDIN

; APPLICANT: SARKAR, FAZLUL H

; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN  
; TITLE OF INVENTION: NEOPLASTIC DISEASE  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG PC  
; STREET: 127 Peachtree Street, Suite 1200  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: usa  
; ZIP: 30303

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/08/229,515A

FILING DATE: 19 APR 1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: PERRYMAN, DAVID G

REGISTRATION NUMBER: 33,438

REFERENCE/DOCKET NUMBER: 1414,608

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404-688-0770

TELEFAX: 404-688-9880

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 4530 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-229-515A-9

## Alignment Scores:

Pred. No.: 4.01e-32 Length: 4530  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

SEQ5 (1-59) x US-08-229-515A-9 (1-4530)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
Db 3118 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTCTACCGCTCACTGCTG 3177  
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
Db 3178 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCACGACGAGGC 3237  
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59  
Db 3238 TTCTTCTGTCAGACCTGCCCGGGCGCTGGGGGCATGGTCCACACGACGACCGC 3294

## RESULT 6

US-08-645-865-9

; Sequence 9, Application US/08645865

; Patent No. 5654406

; GENERAL INFORMATION:

; APPLICANT: RAZIUDIN

; APPLICANT: SARKAR, FAZLUL H

; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN

; TITLE OF INVENTION: NEOPLASTIC DISEASE

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NEEDLE &amp; ROSENBERG PC

; STREET: 127 Peachtree Street, Suite 1200

; CITY: Atlanta

; STATE: Georgia

; COUNTRY: usa

; ZIP: 30303

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-645-865-9

Alignment Scores:
Pred. No.: 4,01e-32 Length: 4530
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

SEQ5 (1-59) x US-08-645-865-9 (1-4530)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3118 CAGAAAGAGGACTTGGGGCCAGCCAGCTGCTGGACAGCACCTTCTACCGCTCAGCTG 3177
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3178 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3237
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
Db 3238 TTCTTCTGTCCAGACCTGCCCCGGCGCTGGGGGCATGCTCCACACAGGCACCGC 3294

RESULT 7
US-09-167-322-4
; Sequence 4, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
```

```
;
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-167-322-4

Alignment Scores:
Pred. No.: 4,01e-32 Length: 4530
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

SEQ5 (1-59) x US-09-167-322-4 (1-4530)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3118 CAGAAAGAGGACTTGGGGCCAGCCAGCTGCTGGACAGCACCTTCTACCGCTCAGCTG 3177
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3178 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3237
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
Db 3238 TTCTTCTGTCCAGACCTGCCCCGGCGCTGGGGGCATGCTCCACACAGGCACCGC 3294

RESULT 8
US-09-527-487-1
; Sequence 1, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(3915)
;
US-09-527-487-1

Alignment Scores:
Pred. No.: 4,01e-32 Length: 4530
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

SEQ5 (1-59) x US-09-527-487-1 (1-4530)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3118 CAGAAAGAGGACTTGGGGCCAGCCAGCTGCTGGACAGCACCTTCTACCGCTCAGCTG 3177
```



```

Db 3118 CAGAATGAGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3177
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 3178 GAGGACGATGACATGGGACCTGTGGTATCTGAGGATATCTGGTACCCAGCAGGCG 3237
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3238 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCGATGGTCCACACAGCAGCAGCG 3294

RESULT 9
US-09-877-177A-11
; Sequence 11, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-11

Alignment Scores:
Pred. No.: 4,01e-32 Length: 4530
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

SEQ5 (1-59) x US-09-877-177A-11 (1-4530)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3118 CAGAATGAGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3177
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 3178 GAGGACGATGACATGGGACCTGTGGTATCTGAGGATATCTGGTACCCAGCAGGCG 3237
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3238 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCGATGGTCCACACAGCAGCAGCG 3294

RESULT 10
US-08-229-515A-14
; Sequence 14, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; 
```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-229-515A-14

Alignment Scores:
Pred. No.: 3,49e-26 Length: 3955
Score: 279.00 Matches: 50
Percent Similarity: 88.14% Conservative: 2
Best Local Similarity: 84.75% Mismatches: 7
Query Match: 85.58% Indels: 0
DB: 1 Gaps: 0

SEQ5 (1-59) x US-08-229-515A-14 (1-3955)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2999 CAGAACGAGGACTGGGGCCATCCAGCCCCAGGACAGTACCTTACCGTCACTGCTG 3058
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 3059 GAAGATGATGACATGGTGACCTGTAGACGCTGAAGAGTATCTGTCGCCACGAGGGA 3118
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
Db 3119 TTCTTCTCCCGGACCCCTACCCAGGCACCTGGGAGCACAGCCCATAGAAGGACCGC 3175

RESULT 11
US-08-645-865-14
; Sequence 14, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIUDDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; 
```

```
;
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-645-865-14

Alignment Scores:
Pred. No.: 3,49e-26 Length: 3955
Score: 279.00 Matches: 50
Percent Similarity: 88.14% Conservative: 2
Best Local Similarity: 84.75% Mismatches: 7
Query Match: 85.58% Indels: 0
DB: 1 Gaps: 0

SEQ5 (1-59) x US-08-645-865-14 (1-3955)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2999 CAGAAGCAGGACTGTGGCCCATCCAGCCCATGAGCAGTACCTTCCACGCTG 3058
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 3059 GAAGATGATGACATGGTGTACCTGGTAGACGCTGAAGAGTATCTGTGCCCCAGCAGGGA 3118
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 3119 TTCTTCCCGGACCCATCCCGAGCAGCTGGGAGCAGACGCCCATAGAAGGCCCGC 3175

RESULT 12
US-08-475-035-3
; Sequence 3, Application US/08475035
; Patent No. 5985553
; GENERAL INFORMATION:
; APPLICANT: KING, C. R.
; APPLICANT: KRAUS, MATTHIAS H.
; APPLICANT: AARONSON, STUART A.
; TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM
; TITLE OF INVENTION: EGF RECEPTOR GENE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, 127 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,035
; FILING DATE: 7 Jun 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perryman, David G.
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.656
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 187...3816
US-08-475-035-3

Alignment Scores:
Pred. No.: 5,46e-08 Length: 5532
Score: 138.00 Matches: 23
Percent Similarity: 91.18% Conservative: 8
Best Local Similarity: 67.65% Mismatches: 3
Query Match: 42.33% Indels: 0
DB: 2 Gaps: 0

SEQ5 (1-59) x US-08-475-035-3 (1-5532)
QY 9 SerProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeu 28
Db 3157 AGTCCTACAGACTCCAACCTTCTACCGTGCCTGATGATGAAGACATGGACGACGTG 3216
QY 29 ValAspAlaGluTyrLeuValProGlnGlnGlyPhePhe 42
Db 3217 GTGGATGCCGACGAGTACCTATCCACACAGCAGGGCTTCTTC 3258

RESULT 13
US-09-676-610B-17
; Sequence 17, Application US/09676610B
; Patent No. 6444465
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Jacqueline Wyatt
; APPLICANT: Susan M. Freiler
; TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HER-1 EXPRESSION
; FILE REFERENCE: RTS-0138
; CURRENT APPLICATION NUMBER: US/09/676,610B
; CURRENT FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 182
; SEQ ID NO 17
; LENGTH: 5532
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (187)...(3819)
US-09-676-610B-17

Alignment Scores:
Pred. No.: 5,46e-08 Length: 5532
Score: 138.00 Matches: 23
Percent Similarity: 91.18% Conservative: 8
Best Local Similarity: 67.65% Mismatches: 3
Query Match: 42.33% Indels: 0
DB: 4 Gaps: 0

SEQ5 (1-59) x US-09-676-610B-17 (1-5532)
QY 9 SerProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeu 28
Db 3157 AGTCCTACAGACTCCAACCTTCTACCGTGCCTGATGATGAAGACATGGACGACGTG 3216
QY 29 ValAspAlaGluTyrLeuValProGlnGlnGlyPhePhe 42
Db 3217 GTGGATGCCGACGAGTACCTATCCACACAGCAGGGCTTCTTC 3258

RESULT 14
US-09-877-177A-10
; Sequence 10, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 56.0384 Seconds  
(without alignments)  
2842.104 Million cell updates/sec

Title: SEQ5

Perfect score: 326

Sequence: 1 QNEDIGPASPDLSTFYRSLL.....GFFCPDPAPGAGMWHHRH 59

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlp  
-O=/cn2.1/USPTO\_spool\_p/HOLLERAN480/runat\_15102003\_131912\_20526/app\_query.fasta\_1.4685  
-DB=N\_Geneseq\_19jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 -CGN 1.1.1758 @runat\_15102003\_131912\_20526 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPELCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -FGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	326	100.0	1115	21	AAF21778	Human breast and o
2	326	100.0	1755	24	AAD32746	Human CDNA for the
3	326	100.0	1767	24	AAD32744	Human CDNA for the
4	326	100.0	1773	24	AAD32747	Human CDNA for the
5	326	100.0	1806	24	AAD32745	Human CDNA for the
6	326	100.0	3600	21	AAA89736	Human HER-2/neu co
7	326	100.0	3678	24	ABK86207	cDNA encoding huma
8	326	100.0	3768	17	AAT40239	HER-2/neu oncogene
9	326	100.0	3768	20	AAK01912	Human HER-2/neu on
10	326	100.0	3768	21	AAA09455	Human heregulin 2
11	326	100.0	3768	22	AAH23392	Human HER-2/neu pr
12	326	100.0	3768	24	ABZ35744	Human ERBB2 polyu
13	326	100.0	3768	24	ABX09987	Human ERBB2 DNA fr
14	326	100.0	3768	24	AAD43935	Human HER-2 cDNA.
15	326	100.0	3768	24	AAD43986	Human HER2 antigen
16	326	100.0	3768	24	ABV78168	Human ERBB2 DNA SE
17	326	100.0	3768	24	AAD32743	Human HER-2/neu pr
18	326	100.0	3768	24	ABA92250	Human HER-2/neu CD
19	326	100.0	3768	24	ABK10730	Human HER-2/neu DN
20	326	100.0	3768	24	ABL91709	Human polynucleoti
21	326	100.0	3768	24	ABK114058	Human HER2 (Erbb2)
22	326	100.0	4299	14	AAQ46083	Sequence encoding
23	326	100.0	4472	21	AAA14812	cDNA encoding the
24	326	100.0	4473	19	ABQ76220	Human tumour anti
25	326	100.0	4473	20	AAZ31071	HER-2 nucleic acid
26	326	100.0	4473	24	ABZ34969	Human gene express
27	326	100.0	4473	24	ABZ38904	Human Her-2 DNA.
28	326	100.0	4530	16	AAT01585	Her-2/neu (ERBB2/c
29	326	100.0	4530	18	AAT71253	Human HER2 gene.
30	326	100.0	4530	21	AAZ60815	Nucleotide sequenc
31	326	100.0	4530	22	AAD19731	Human tyrosine kin
32	326	100.0	4530	24	ABZ35012	Human gene express
33	326	100.0	4530	24	ABV94128	Breast carcinoma r
34	326	100.0	4530	24	ABN85585	Human HER2-neu SEQ
35	326	100.0	4530	24	ABK83918	Human CDNA differe
36	326	100.0	4530	25	ACC50139	Breast cancer asso
37	326	100.0	4530	25	ABQ83856	Human Her2/Neu enc
38	326	100.0	9274	22	AAF24297	HER2 transgene pla
39	326	100.0	9274	24	AAD43934	HER-2 transgene pl
40	326	100.0	9274	24	ABK14057	Human HER2 (Erbb2)
41	279	85.6	3955	16	AAT01590	Rat neu promoter.
42	279	85.6	3955	21	AAA89753	Rat HER-2/neu prot
43	269	82.5	2763	24	ABA92252	Mouse Her-2/neu ex
44	269	82.5	2781	24	ABA92253	Her-2/neu extracel
45	269	82.5	3771	21	AAA89737	Mouse Her-2/neu CD

ALIGNMENTS

RESULT 1

AAF21778

ID AAF21778 standard; DNA; 1115 BP.

XX

AC AAF21778;

XX 27-MAR-2001 (first entry)

DE Human breast and ovarian cancer associated antigen gene SEQ ID 165.

XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;  
KW nontropic; neuroprotective; antiviral; antiallergic; hepatotropic;  
KW antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;  
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;  
KW Addison's disease; allergy; autoimmune haemolytic anaemia;  
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;

KW cardiovascular disorder; wound healing; neurological disease; ds.  
 XX Homo sapiens.  
 XX W0200055173-A1.  
 PN 21-SEP-2000.  
 XX 08-MAR-2000; 2000WO-US05881.  
 XX 12-MAR-1999; 99US-0124270.  
 PR (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Ruben SM;  
 XX WPI; 2000-611515/58.  
 DR P-PSDB; AAB58875.  
 XX New human breast and ovarian cancer associated gene sequences and the  
 PT polypeptides encoded by these genes, useful in the prevention,  
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular  
 PT disorders and neurological diseases  
 XX Claim 1; Page 604; 1299pp; English.  
 PS  
 XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are  
 CC associated with breast and ovarian cancer. Included in the invention are  
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
 CC isolation and characterisation of the DNA and protein sequences of the  
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist  
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;  
 CC neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;  
 CC antidiabetic; antiinflammatory; antitumor; antitumor; anticonvulsant;  
 CC antibacterial; antifungal; antiparasitic and cardiant activity. The  
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,  
 CC particularly breast and ovarian cancer. The nucleic acid sequences,  
 CC proteins, agonists and antagonists may also be used in the diagnosis,  
 CC prevention and treatment of immune disorders e.g. Addison's disease,  
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
 CC arthritis and ulcerative colitis; cardiovascular disorders such as  
 CC myocardial ischaemias; wound healing; neurological diseases such as  
 CC cerebral anoxia and epilepsy; and infectious diseases.  
 XX  
 XX Sequence 1115 BP; 210 A; 336 C; 338 G; 222 T; 9 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1.4e-32 Length: 1115  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 21 Gaps: 0  
 SEQ5 (1-59) x AAF21778 (1-1115)  
 Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTrArgSerLeuLeu 20  
 Db 455 CAGNATGAGGACTTGGGCCCCAGCCAGTCCCTTGACAGACACCTTCTACCGCTCAGTCTG 514  
 Qy 21 GlnAspAspMetGlyAspLeuValAspAlaGluGlyTrLeuValProGlnGlnGly 40  
 Db 515 GAGGACCATGACATGGGGGACCTGGTGATGCTGAGAGATATCTGGTACCCAGCAGGC 574  
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59  
 Db 575 TTCTCTGTGTGACAGACCTGCCCGGGGCTGGGGGCGATGTCACACACAGCAGCCGC 631  
 RESULT 2  
 AAD32746  
 ID AAD32746 standard; cDNA; 1755 BP.

XX AAD32746;  
 AC 01-JUL-2002 (first entry)  
 XX Human cDNA for the clone HICD\_native\_coding\_region.  
 DE Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;  
 XX human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.  
 KW Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 AC 1..1755  
 FT /tag= a  
 FT /product= "Human protein encoded by cDNA for the clone  
 FT HICD\_native\_coding\_region"  
 FT /transl\_except= (pos:1741..1752, aa:Leu-Glu)  
 FT /note= "CDS does not include stop codon"  
 FT /partial  
 XX W0200214503-A2.  
 XX 21-FEB-2002.  
 PD 14-AUG-2001; 2001WO-US41733.  
 XX 14-AUG-2000; 2000US-225152P.  
 PR 28-SEP-2000; 2000US-236428P.  
 PR 21-FEB-2001; 2001US-270520P.  
 XX (CORI-) CORIXA CORP.  
 XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kaloş MD;  
 PI McNeill PD, Vedvick TS;  
 PI WPI; 2002-280758/32.  
 DR P-PSDB; AAE20483.  
 XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,  
 XX prevention and diagnosis of cancer, preferably breast cancer  
 XX Example 5; Page 118-119; 129pp; English.  
 CC The invention relates to an isolated Her-2/Neu polypeptide composition  
 CC effective for eliciting an immune response. The invention is useful for  
 CC eliciting an immune response in a patient, where the patient is human  
 CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.  
 CC The composition is useful for the therapy and diagnosis of cancer,  
 CC and other compositions for the diagnosis, prevention and treatment of  
 CC human malignancies, for stimulating and/or expanding T cells specific for  
 CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a  
 CC patient. The invention is useful for stimulating a T cell response in a  
 CC human patient, as probe or primer for nucleic acid hybridisation, to  
 CC selectively form duplex molecules with complementary stretches of the  
 CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full  
 CC length gene from a suitable library, and to direct expression of a  
 CC polypeptide in appropriate host cells. The composition is useful in  
 CC prophylactic or therapeutic applications and for the treatment of cancer,  
 CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-  
 CC associated malignancies. The invention is useful in gene therapy. The  
 CC present sequence is human cDNA for the clone HICD\_native\_coding\_region.  
 XX  
 SQ Sequence 1755 BP; 376 A; 517 C; 531 G; 331 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2.4e-32 Length: 1755  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

```

SEQ5 (1-59) x AAD32746 (1-1755)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
    |||||
Db 946 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1005
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
    |||||
Db 1006 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGTACCCAGCAGGCG 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
    |||||
Db 1066 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCGATGGTCCACACAGGACGCGC 1122

RESULT 3
AAD32744
ID AAD32744 standard; cDNA; 1767 BP.
AC AAD32744;
DT 01-JUL-2002 (first entry)
DE Human cDNA for the clone HICD_CT_His_coding_region.
KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1764
FT /tag= a
FT /product= "Human protein encoded by cDNA for the clone
FT HICD_CT_His_coding_region"
XX
PN WO200214503-A2.
XX
PD 21-FEB-2002.
XX
PF 14-AUG-2001; 2001WO-US41733.
XX
PR 14-AUG-2000; 2000US-225152P.
PR 28-SEP-2000; 2000US-236428P.
PR 21-FEB-2001; 2001US-270520P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
PI McNeill PD, Vedvick TS;
XX
XX WPI; 2002-280758/32.
XX P-PSDB; AAE20481.
XX
XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,
XX prevention and diagnosis of cancer, preferably breast cancer
XX
XX Example 5; Page 117-118; 129pp; English.
XX
XX The invention relates to an isolated Her-2/Neu polypeptide composition
XX effective for eliciting an immune response. The invention is useful for
XX eliciting an immune response in a patient, where the patient is human
XX leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
XX The composition is useful for the therapy and diagnosis of cancer,
XX preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
XX and other compositions for the diagnosis, prevention and treatment of
XX human malignancies, for stimulating and/or expanding T cells specific for
XX Her-2/Neu polypeptide and for inhibiting the development of cancer in a
XX patient. The invention is useful for stimulating a T cell response in a
XX human patient, as probe or primer for nucleic acid hybridisation, to
XX selectively form duplex molecules with complementary stretches of the
XX entire Her-2/Neu gene or gene fragments of interest, to isolate a full
XX length gene from a suitable library, and to direct expression of a

```

```

CC polypeptide in appropriate host cells. The composition is useful in
CC prophylactic or therapeutic applications and for the treatment of cancer,
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
CC associated malignancies. The invention is useful in gene therapy. The
CC present sequence is human cDNA for the clone HICD_CT_His_coding_region.
XX
SQ Sequence 1767 BP; 381 A; 521 C; 529 G; 336 T; 0 other;

Alignment Scores:
Pred. No.: 2,42e-32 Length: 1767
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

SEQ5 (1-59) x AAD32744 (1-1767)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
    |||||
Db 946 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1005
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
    |||||
Db 1006 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGTACCCAGCAGGCG 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59
    |||||
Db 1066 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCGATGGTCCACACAGGACGCGC 1122

RESULT 4
AAD32747
ID AAD32747 standard; cDNA; 1773 BP.
XX
AC AAD32747;
XX
DT 01-JUL-2002 (first entry)
XX
DE Human cDNA for the clone HICD_in_ppDM_coding_sequence.
XX
KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT CDS 1..1770
XX FT /tag= a
XX FT /product= "Human protein encoded by cDNA for the clone
XX HICD_in_ppDM_coding_sequence"
XX
XX WO200214503-A2.
XX
XX 21-FEB-2002.
XX
XX 14-AUG-2001; 2001WO-US41733.
XX
XX 14-AUG-2000; 2000US-225152P.
XX 28-SEP-2000; 2000US-236428P.
XX 21-FEB-2001; 2001US-270520P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
PI McNeill PD, Vedvick TS;
XX
XX WPI; 2002-280758/32.
XX P-PSDB; AAE20484.
XX
XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,
XX prevention and diagnosis of cancer, preferably breast cancer
XX
XX Example 5; Page 119; 129pp; English.
XX

```

```

XX The invention relates to an isolated Her-2/Neu polypeptide composition
CC effective for eliciting an immune response. The invention is useful for
CC eliciting an immune response in a patient, where the patient is human
CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
CC The composition is useful for the therapy and diagnosis of cancer,
CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
CC and other compositions for stimulating and/or expanding T cells specific for
CC human malignancies, for stimulating and/or inhibiting the development of a
CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
CC patient. The invention is useful for stimulating a T cell response in a
CC human patient, as probe or primer for nucleic acid hybridisation, to
CC selectively form duplex molecules with complementary stretches of the
CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
CC length gene from a suitable library, and to direct expression of a
CC polypeptide in appropriate host cells. The composition is useful in
CC prophylactic or therapeutic applications and for the treatment of cancer,
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
CC associated malignancies. The invention is useful in gene therapy. The
CC present sequence is human cDNA for the clone HICD_in_ppdm_coding_
XX sequence.
SQ Sequence 1773 BP; 383 A; 528 C; 530 G; 332 T; 0 other;

Alignment Scores:
Pred. No.: 2,43e-32 Length: 1773
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

SEQ5 (1-59) x AAD32747 (1-1773)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20
DB 970 CAGAAATGAGGACTTGGGCCCGAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTGTG 1029

QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
DB 1030 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGAGTATCTGTACCCAGCAGGGC 1089

QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
DB 1090 TTCTTGTGTGACAGACCTGCCCGGGCGTGGGGGATGCTGCCACACAGCAGCCG 1146

RESULT 5
AAD32745
ID AAD32745 standard; cDNA; 1806 BP.
XX
AC AAD32745;
XX
XX 01-JUL-2002 (first entry)
XX
DE Human cDNA for the clone HICD_plus_8_HIS.
XX
XX Human; Her-2/Neu protein; immune response; gene therapy: breast cancer;
KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 1..1803
FT /*tag= a
FT /product= "Human protein encoded by cDNA for the clone
FT Hicb_plus_8_HIS"
FT /transl_except= (pos:1543..1545, aa:Pro)
XX
XX WO200214503-A2.
XX
XX 21-FEB-2002.
XX
XX 14-AUG-2001; 2001WO-US41733.

```

```

XX 14-AUG-2000; 2000US-225152P.
PR 28-SEP-2000; 2000US-236428P.
PR 21-FEB-2001; 2001US-270520P.
XX (CORI-) CORIXA CORP.
XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
PI McNeill PD, Vedvick TS;
XX WPI; 2002-280758/32.
DR P-PSDB; AAE20482.
XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,
PT prevention and diagnosis of cancer, preferably breast cancer -
XX Example 5; Page 118; 129pp; English.
XX The invention relates to an isolated Her-2/Neu polypeptide composition
CC effective for eliciting an immune response. The invention is useful for
CC eliciting an immune response in a patient, where the patient is human
CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
CC The composition is useful for the therapy and diagnosis of cancer,
CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
CC and other compositions for the diagnosis, prevention and treatment of
CC human malignancies, for stimulating and/or expanding T cells specific for
CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
CC patient. The invention is useful for stimulating a T cell response in a
CC human patient, as probe or primer for nucleic acid hybridisation, to
CC selectively form duplex molecules with complementary stretches of the
CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
CC length gene from a suitable library, and to direct expression of a
CC polypeptide in appropriate host cells. The composition is useful in
CC prophylactic or therapeutic applications and for the treatment of cancer,
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
CC associated malignancies. The invention is useful in gene therapy. The
CC present sequence is human cDNA for the clone HICD_plus_8_HIS.
XX Sequence 1806 BP; 391 A; 530 C; 544 G; 341 T; 0 other;

Alignment Scores:
Pred. No.: 2,48e-32 Length: 1806
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

SEQ5 (1-59) x AAD32745 (1-1806)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20
DB 1003 CAGAAATGAGGACTTGGGCCCGAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTGTG 1062

QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
DB 1063 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGAGTATCTGTACCCAGCAGGGC 1122

QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
DB 1123 TTCTTGTGTGACAGACCTGCCCGGGCGTGGGGGATGCTGCCACACAGCAGCCG 1179

RESULT 6
AAD32745
ID AAA89736 standard; DNA; 3600 BP.
XX
XX AAA89736;
XX
XX 12-JAN-2001 (first entry)
XX
XX Human HER-2/neu coding sequence.
XX
XX Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
KW

```





```

Db      2878 CAGAAATGAGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTACCGCTCAGTCTG 2937
QY      21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
      |||||
Db      2938 GAGACGATGACATGGGGACCTTGGTGATCTCTGAGAGTATCTGGTACCCAGCAGGC 2997
QY      41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
      |||||
Db      2998 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGTGTCCACCAGCAGCCGC 3054

RESULT 8
AAT40739
ID      AAT40739 standard; cDNA; 3768 BP.
XX
AC      AAT40739;
XX
DT      01-JAN-1997 (first entry)
XX
DE      HER-2/neu oncogene.
XX
KW      HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;
KW      breast cancer; ovary cancer; colon cancer; lung cancer;
KW      prostate cancer; genetic immunisation; tumour; vaccine; vector;
KW      ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      CDS
      1..3765
      /*tag= b
      /note= "nucleotides 2026-3765 (claim 1) code for
      HER-2/neu intracellular domain"
XX
PN      WO9630514-A1.
XX
PD      03-OCT-1996.
XX
PF      28-MAR-1996; 96WO-US01689.
XX
PR      31-MAR-1995; 95US-0414417.
XX
PA      (UNIW ) UNIV WASHINGTON.
XX
PI      Cheever MA, Disis ML;
XX
WPI: 1996-455361/45.
P-PSDB: AAW01111.
XX
DNA encoding HER-2/neu polypeptide(s) - used for prevention or
treatment of malignancies with which the HER-2/neu oncogene is
associated
XX
Claim 1; Page 49-56; 7lpp; English.
XX
Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or
c-erbB2) protein (AAW01111). The oncogene is overexpressed in various
cancers, including breast, ovarian, colon, lung and prostate, and
appears to induce malignancies through quantitative mechanisms that
result from increased or deregulated expression of an essentially
normal gene product. Nucleotides 2026-3765 of the cDNA sequence
code for the intracellular domain (tyrs676-val1255) of the HER-2/neu
protein, which is useful for immunisation against malignancy.
XX
Nucleic acids can be used to direct expression of the intracellular
domain in transformed host cells, or are used, alone or in a viral
vector, for genetic immunisation of an animal.
XX
SQ      Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 5 97e-32 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```

```

Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0
SEQ5 (1-59) x AAT40739 (1-3768)
QY      1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
      |||||
Db      2968 CAGAATGAGACCTTGGGCCCGCCAGTCCTTGGACAGCACCTTCTACCGCTCAGTCTG 3027
QY      21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
      |||||
Db      3028 GAGACGATGACATGGGGACCTTGGTGATCTCTGAGAGTATCTGGTACCCAGCAGGC 3087
QY      41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
      |||||
Db      3088 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGTGTCCACCAGCAGCCGC 3144

RESULT 9
AAX01912
ID      AAX01912 standard; DNA; 3768 BP.
XX
AC      AAX01912;
XX
DT      21-APR-1999 (first entry)
XX
DE      Human HER-2/neu oncogene DNA.
XX
KW      HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
KW      malignancy; treatment; tumour; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      CDS
      1..3768
      /*tag= a
      /product= "HER-2/neu"
      /note= "Oncogene"
      2026..3765
      /*tag= b
      /note= "region which elicits immune response"
XX
PN      US5869445-A.
XX
PD      09-FEB-1999.
XX
PF      01-APR-1996; 96US-0625101.
XX
PR      01-APR-1996; 96US-0625101.
PR      17-MAR-1993; 93US-0033644.
PR      12-AUG-1993; 93US-0106112.
PR      31-MAR-1995; 95US-0414417.
XX
PA      (UNIW ) UNIV WASHINGTON.
XX
PI      Cheever MA, Disis ML;
XX
WPI: 1999-152835/13.
P-PSDB: AAW92406.
XX
PT      Use of HER-2/neu polypeptides - for eliciting an immune response to
PT      an HER-2/neu associated malignancy, particularly for treating or
PT      preventing tumours
XX
PS      Claim 1a; Column 23-32; 26pp; English.
XX
This sequence encodes the human HER-2/neu oncogene protein. A fragment
of this protein is used in a method for eliciting or enhancing an immune
response to HER-2/neu protein. The polypeptide can stimulate T cells and
B cells to produce an immune response to the HER-2/neu protein. The
method can be used for immunisation against a malignancy in which the
HER-2/neu oncogene is associated and in the treatment of an existing
tumour, or to prevent tumour occurrence or reoccurrence.
XX

```

SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:  
 Pred. No.: 5,97e-32 Length: 3768  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0

SEQ5 (1-59) x AAX01912 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 Db 2968 CAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027  
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
 Db 3028 GAGGACGATGACATGGGGGACCTGCTGTGATGCTGAGGAGTATCTGGTACCCAGCAGCGGC 3087  
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59  
 Db 3088 TTCTCTCTCCAGACCTGCCCGCGCTGGGGCGATGGTCCACCAGCAGCAGCGC 3144

#### RESULT 10

AAA09455  
 ID AAA09455 standard; DNA; 3768 BP.  
 AC AAA09455;  
 XX

DT 10-AUG-2000 (first entry)

XX Human heregulin 2 (Her2) coding sequence.

XX Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;  
 KW self-protein; cell-associated peptide antigen; foreign epitope;  
 KW cancer; breast cancer; prostate cancer; ss.

XX Homo sapiens.

OS WO200020027-A2.

PN 13-APR-2000.

XX 05-OCT-1999; 99WO-DK00525.

PF 05-OCT-1998; 98DK-0001261.

PR 20-OCT-1998; 98US-0105011.

XX (MEBI-) M & E BIOTECH AS.

PA Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;  
 PI Gautam A, Birk P, Karlsson G;

XX WPI; 2000-349917/30.

DR P-PSDB; AAY92620.

XX Inducing immune responses to weakly immunogenic, tumor associated  
 PT peptide antigens for the treatment of breast and prostate cancer

XX Claim 62; Page 187-193; 220pp; English.

XX The claims detail a method for inducing immune responses against weakly  
 CC immunogenic cell-associated peptide antigens (PA) such as those  
 CC associated with cancers (i.e. self-proteins), for example, human  
 CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or  
 CC fibroblast growth factor 8b (FGF8b). The method comprises effecting  
 CC simultaneous presentation by antigen producing cells (APCs) of the  
 CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)  
 CC group derived from the PA and/or at least 1 B-cell group derived from the  
 CC cell-associated PA; and (2) at least 1 first T helper cell group which is  
 CC foreign to the animal. Analogues of human PSM, human Her2 and  
 CC human/murine FGF8b comprising a substantial part of all known and

CC predicted CTL and B-cell epitopes of the respective PA and including at  
 CC least one foreign T helper epitope are also claimed. The method is used  
 CC to treat prostate, prostate/breast or breast cancer when the PA is human  
 CC PSM, FGF8b and Her2, respectively.

XX SQ Sequence 3768 BP; 759 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:  
 Pred. No.: 5,97e-32 Length: 3768  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 21 Gaps: 0

SEQ5 (1-59) x AAA09455 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 Db 2968 CAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027  
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
 Db 3028 GAGGACGATGACATGGGGGACCTGCTGTGATGCTGAGGAGTATCTGGTACCCAGCAGCGC 3087  
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 59  
 Db 3088 TTCTCTCTCCAGACCTGCCCGCGCTGGGGCGATGGTCCACCAGCAGCAGCGC 3144

#### RESULT 11

AAH23392

ID AAH23392 standard; DNA; 3768 BP.

XX AC AAH23392;

XX DT 25-SEP-2001 (first entry)

XX Human HER-2/neu protein encoding DNA.

XX Antigen-presenting cell; immunogenic; immune response; HER-2/neu;  
 KW oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.

XX Homo sapiens.

OS Key Location/Qualifiers

FT CDS 1..3768

FT /\*tag= a

FT /product= "HER-2/neu protein"

XX WO200153463-A2.

XX 26-JUL-2001.

XX 19-JAN-2001; 2001WO-US01850.

XX 21-JAN-2000; 2000US-0177545.

XX (CORI-) CORIXA CORP.

XX Cheever MA, Hand-Zimmermann S;

XX WPI; 2001-476112/51.

DR P-PSDB; AAB85458.

XX New antigen-presenting cells, useful as vaccines for eliciting or  
 PT enhancing an immune response to HER-2/neu protein, particularly useful  
 PT for treating or preventing cancer, e.g. breast cancer

PS Claim 1; Page 41-46; 49pp; English.

XX The invention provides an isolated antigen-presenting cell, which  
 CC expresses at least an immunogenic portion of a polypeptide that produces  
 CC an immune response to HER-2/neu protein. The antigen-presenting cells are

CC useful as vaccines for eliciting or enhancing an immune response to  
 CC HER-2/neu protein, particularly in treating or preventing malignancies in  
 CC which the HER-2/neu oncogene is associated. Specifically, these are  
 CC useful for treating or preventing cancer, e.g. breast cancer, ovarian,  
 CC colon, lung or prostate cancers. The present sequence represents a DNA  
 CC encoding the human HER-2/neu protein (also known as p185 or c-erbB2).  
 XX

SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:  
 Pred. No.: 5.97e-32 Length: 3768  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0

SEQ5 (1-59) x AAH23392 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 |||||  
 Db 2968 CAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACGTGCTG 3027

QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40  
 |||||  
 Db 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTATACCCAGCAGGC 3087

QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59  
 |||||  
 Db 3088 TTCTTCTGTCCAGACCTGCCCCGGGGCGTGGGGGCGATGTCACACAGGCACCGC 3144

RESULT 12  
 ABZ35744

ID ABZ35744 standard; DNA; 3768 BP.

AC ABZ35744;

XX 07-FEB-2003 (first entry)

XX Human ERBB2 polynucleotide SEQ ID NO 52.

DE Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;  
 KW protozoacide; gene expression; antisense; tumour; infection; plasmodium;  
 KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;  
 KW Hepatitis C virus; human papilloma virus; gene; ds.

XX Homo sapiens.

XX DE10100588-Al.

XX 18-JUL-2002.

XX 09-JAN-2001; 2001DE-1000588.

XX 09-JAN-2001; 2001DE-1000588.

XX (RIBO-) RIBOPHARMA AG.

XX Kreutzer R, Limmer S, Rost S, Hadwiger P;

XX WPI; 2002-683450/74.

XX Inhibiting expression of target genes, useful e.g. for treating tumors,  
 PT by introducing into cells two double-stranded RNAs that are  
 PT complementary to the target

PS Claim 13; Page 38-39; 100pp; German.

XX The invention relates to inhibiting expression of a target gene in a cell  
 CC by introducing at least two oligoribonucleotides (dsRNAi and II), both  
 CC with a double-stranded (ds) structure of at most 49 sequential nucleotide  
 CC pairs. At least part of one strand (S1, S2) of the ds structures in each  
 CC of dsRNAi and II are complementary to regions in the target gene. The

CC method uses antisense inhibition of gene expression using double stranded  
 CC RNA inhibition (RNAi). The method is particularly used to treat tumours  
 CC or infections, especially by plasmodium or viruses/viroids (pathogenic on  
 CC humans, animals or plants). The method provides more effective inhibition  
 CC of expression than known methods using a single dsRNA, even at very low  
 CC concentrations. When dsRNA has at least one unpaired nucleotide at the  
 CC end, stability (and thus effective concentration in the cell) is  
 CC improved and efficiency can be increased further by pretreating the cells  
 CC with interferon. The present sequence is that of a target DNA of the  
 CC invention.

XX SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:  
 Pred. No.: 5.97e-32 Length: 3768  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

SEQ5 (1-59) x ABZ35744 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 |||||  
 Db 2968 CAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACGTGCTG 3027

QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40  
 |||||  
 Db 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTATACCCAGCAGGC 3087

QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59  
 |||||  
 Db 3088 TTCTTCTGTCCAGACCTGCCCCGGGGCGTGGGGGCGATGTCACACAGGCACCGC 3144

RESULT 13

ABX09987

ID ABX09987 standard; DNA; 3768 BP.

XX AC ABX09987;

XX 23-JAN-2003 (first entry)

XX Human ERBB2 DNA fragment SEQ ID 52.

DE Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;  
 KW prion; inhibition; human; ds.

XX Homo sapiens.

XX DE10100587-C1.

XX 21-NOV-2002.

XX 09-JAN-2001; 2001DE-1000587.

XX 09-JAN-2001; 2001DE-1000587.

XX (RIBO-) RIBOPHARMA AG.

XX Kreutzer R, Limmer S, Rost S, Hadwiger P;

XX WPI; 2002-742209/81.

XX Inhibiting expression of target genes, e.g. oncogenes, in cells, by  
 PT introduction of complementary double-stranded oligoribonucleotide,  
 PT after treating the cell with interferon

XX Disclosure; Page 43-44; 98pp; German.

XX This invention describes a novel method for inhibiting expression of a  
 CC target gene by introducing into the cell that contains the target gene  
 CC at least one oligoribonucleotide (dsRNAi) that has a double-stranded

CC (ds) structure of not more than 49 consecutive nucleotides (nt), where  
 CC at least a segment of one strand of the ds structure is complementary  
 CC with the target gene and the cells are treated with interferon before  
 CC introduction of dsRNA1. The method is used to inhibit expression of  
 CC target genes, particularly oncogenes, cytokine genes, Id (not defined)  
 CC protein genes; developmental or prion genes, or genes expressed in  
 CC pathogenic organisms (particularly plasmodia) or in viruses or viroids  
 CC (pathogenic in humans, animals or plants). Treating the cells with  
 CC interferon greatly increases the extent to which dsRNA can inhibit  
 CC expression of the target genes, and the effect is even greater when dsRNA  
 CC are modified to increase their stability. ABX09936-ABX10075 represent  
 CC gene fragments used to illustrate the method of the invention.

XX Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:  
 Pred. No.: 5,97e-32 Length: 3768  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

SEQ5 (1-59) x ABX09987 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027

QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
 Db 3028 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 3087

QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59  
 Db 3088 TTCCTCTCCAGACCTGCCCCGGCGCTGGGGGCGATGGTCCACACAGCAGCCGC 3144

RESULT 14  
 AAD43935

ID AAD43935 standard; cdna; 3768 BP.

XX AAD43935;

DT 13-DEC-2002 (first entry)

DE Human HER-2 cdna.

XX Transgenic animal; transgenic; mammary gland cell; HER2; tumour;  
 KW cancer; therapy; apoptosis; cytostatic; human; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..3768  
 FT /\*tag= a  
 FT /product= "Human HER2 protein"

XX US200203736-A1.

XX 21-MAR-2002.

XX 16-MAR-2001; 2001US-0811115.

XX 16-MAR-2000; 2000US-189844P.

XX (ERIC/) ERICKSON S.

PA (KING/) KING K.

PA (SCHW/) SCHWALL R.

XX Erickson S, King K, Schwall R;

XX WPI; 2002-401155/43.

DR P-PSDB; AAE26349.

XX New transgenic non-human mammal that produces detectable levels of a  
 PT native human HER2 protein in its mammary gland cells, useful as tumor  
 PT models for testing HER2-directed cancer therapies, and for identifying  
 PT anticancer agents

PS Example 2; Page 24-26; 83pp; English.

XX The invention relates to a transgenic non-human mammal that produces in  
 CC its mammary gland cells detectable levels of a native human HER2 protein  
 CC or its fragment. The transgenic animals are useful as tumour models for  
 CC testing HER2-directed cancer therapies, and for identifying anticancer  
 CC agents. The animals may also be used as source of cells which can be  
 CC immortalised in culture, in screening for compounds that have potential  
 CC as prophylactic or therapeutic treatments of diseases or disorders  
 CC involving expression of HER2. The anti-cancer molecules are useful for  
 CC inducing apoptosis or cell death of cancer cells. The present sequence  
 CC is human HER-2 cdna.

XX Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:  
 Pred. No.: 5,97e-32 Length: 3768  
 Score: 326.00 Matches: 59  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

SEQ5 (1-59) x AAD43935 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027

QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
 Db 3028 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 3087

QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59  
 Db 3088 TTCCTCTCCAGACCTGCCCCGGCGCTGGGGGCGATGGTCCACACAGCAGCCGC 3144

RESULT 15

AAD43986

ID AAD43986 standard; DNA; 3768 BP.

XX AAD43986;

DT 13-DEC-2002 (first entry)

DE Human Her2 antigen DNA.

XX Human; immune response; T-helper cell epitope; chitosan; CTL response;  
 KW vaccine; prostate cancer; breast cancer; Her2 antigen; cytostatic;  
 KW immunostimulant; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..3768  
 FT /\*tag= a  
 FT /product= "Human Her2 antigen"  
 FT stg\_peptide 1..69  
 FT /\*tag= b  
 FT mat\_peptide 70..2250  
 FT /\*tag= c  
 FT /product= "Mature human Her2 antigen"

XX WO200234287-A2.

XX 02-MAY-2002.

XX

```

PF 26-OCT-2001; 2001WO-DK00705.
XX
PR 27-OCT-2000; 2000DK-0001506.
PR 03-NOV-2000; 2000US-245156P.
PR 18-JUN-2001; 2001DK-0000936.
XX
PA (PHAR-) PHARMEXA AS.
XX
XX Beier AM, Gautam A, Mouritsen S;
PI
XX
DR WPI; 2002-463339/49.
DR P-PSDB; AAE26366.
XX
XX Inducing or enhancing an immune response against an antigen,
PT particularly cytotoxic T-lymphocyte responses, for treating or
PT ameliorating prostate or breast cancer, comprises administering the
PT antigen formulated with chitosan -
XX
XX Claim 28; Page 85-90; 97pp; English.
PS
XX The invention relates to a method for inducing or enhancing an immune
CC response against a polypeptide antigen in an animal, including human.
CC The method comprises administering the polypeptide antigen or at least
CC one variant which includes at least one first T-helper cell epitope that
CC is foreign to the animal (foreign TH epitope) and is formulated with
CC chitosan. The polypeptide antigen is weakly immunogenic or non-
CC immunogenic. The invention is used as vaccine. The chitosan and
CC polypeptide antigen or its variant are useful in the preparation of an
CC immunogenic composition for inducing or enhancing an immune response,
CC particularly CTL response, against the polypeptide or protein antigen.
CC The method for inducing or enhancing an immune response is useful in
CC treating or ameliorating cancer, e.g. prostate or breast cancer. The
CC present sequence is human Her2 antigen DNA.
XX
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 5.97e-32 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

SEQ5 (1-59) x AAD43986 (1-3768)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAATGAGGACITGGGCCCGCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTGTG 3027
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGGACCTGGTGGATGCTCAGGAGTATCTGGTACCCAGCAGGGC 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 3088 TTCTTCTGTCCAGACCTGCCCGCGCGCTGGGGCGCATGGTCCACCACAGGCACCGC 3144

```

Search completed: October 15, 2003, 23:50:11  
Job time : 61.0384 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.  
  
OM protein - nucleic search, using frame\_plus\_p2n model  
  
Run on: October 15, 2003, 21:09:13 ; Search time 849.651 Seconds  
(without alignments)  
2840.777 Million cell updates/sec  
  
Title: SEQ5  
Perfect score: 326  
Sequence: 1 QNEDLGASPLDSTFYRSL.....GFFCPDPAPGAGGMVHRRH 59  
  
Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0  
  
Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000  
  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO.spool\_p/HOLLERAN480/runat\_15102003\_131912\_20536/app\_query.fasta\_1.4685  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480.qcgn\_1.1\_2724.@runat\_15102003\_131912\_20536 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.ov:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*

29: em.vi:\*  
30: em.htg\_hum:\*  
31: em.htg\_inv:\*  
32: em.htg\_other:\*  
33: em.htg\_mus:\*  
34: em.htg\_pln:\*  
35: em.htg\_rod:\*  
36: em.htg\_mam:\*  
37: em.htg\_vrt:\*  
38: em.sy:\*  
39: em.htgo\_hum:\*  
40: em.htgo\_mus:\*  
41: em.htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	326	100.0	1755	6	AX384609 Sequence
2	326	100.0	1767	6	AX384607 Sequence
3	326	100.0	1773	6	AX384610 Sequence
4	326	100.0	1806	6	AX384608 Sequence
5	326	100.0	3678	6	AX505114 Sequence
6	326	100.0	3768	6	AR034479 Sequence
7	326	100.0	3768	6	AX060704 Sequence
8	326	100.0	3768	6	AX201817 Sequence
9	326	100.0	3768	6	AX380923 Sequence
10	326	100.0	3768	6	AX384604 Sequence
11	326	100.0	3768	6	AX465456 Sequence
12	326	100.0	3768	6	AX467229 Sequence
13	326	100.0	3768	6	AX481438 Sequence
14	326	100.0	4473	6	AR080259 Sequence
15	326	100.0	4473	6	AR167390 Sequence
16	326	100.0	4473	9	HSERB2R
17	326	100.0	4530	6	AR202597 Sequence
18	326	100.0	4530	6	AR283481 Sequence
19	326	100.0	4530	6	AX282577 Sequence
20	326	100.0	4530	6	AX587649 Sequence
21	326	100.0	4530	6	AX644071 Sequence
22	326	100.0	4530	6	BD005474 Cellular
23	326	100.0	4530	6	I21124 Sequence 9
24	326	100.0	4530	6	I59745 Sequence 9
25	326	100.0	4530	9	HUMHER2A
26	326	100.0	9274	6	AX060703 Sequence
27	326	100.0	13450	9	AB096612 Homo sapi
28	326	100.0	16572	9	AB096614 Homo sapi
29	326	100.0	20271	9	AB096613 Homo sapi
30	326	100.0	30837	9	AY208911 Homo sapi
31	326	100.0	161815	9	AC079199 Homo sapi
32	326	100.0	168585	9	AC040933 Homo sapi
33	326	100.0	198008	2	AC142197 Homo sapi
34	300	92.0	3780	4	AB008451 Canis fam
35	288	88.3	4062	10	D16295 Mesocricetu
36	279	85.6	3955	6	AX380924 Sequence
37	279	85.6	3955	6	I21129 Sequence 14
38	279	85.6	3955	6	I59750 Sequence 14
39	279	85.6	3955	10	X01362 Rat mRNA fo
40	279	85.6	4727	10	AY116182 Rattus no
41	274	84.0	239732	2	AC098491 Rattus no
42	269	82.5	2083	10	BC027080 Mus muscu
43	269	82.5	2763	6	AX380942 Sequence
44	269	82.5	2781	6	AX380944 Sequence
45	269	82.5	3771	6	AX189662 Sequence

ALIGNMENTS

RESULT 1

```

AX384609
LOCUS AX384609 1755 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 6 from Patent WO0214503.
ACCESSION AX384609
VERSION AX384609.1 GI:19577810
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .1755
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 376 a 517 c 531 g 331 t
ORIGIN

Alignment Scores:
Pred. No.: 1.9e-31 Length: 1755
Score: 326.00 Matches: 59
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0

SEQ5 (1-59) x AX384609 (1-1755)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 946 CAGAATGAGGACTTGGCCCGAGCCAGTCCCTTGGACAGCAGCTTCTACCGCTCAGTGCTG 1005
|||||
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 1006 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGTATACCCAGCAGGC 1065
|||||
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
Db 1066 TTCTTCTGTCAGACCTGCCCGGGGCTGGGGGATGTTGTCACACAGCAGCCGC 1122
|||||

RESULT 3
AX384610
LOCUS AX384610 1773 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 7 from Patent WO0214503.
ACCESSION AX384610
VERSION AX384610.1 GI:19577811
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .1773
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 383 a 528 c 530 g 332 t
ORIGIN

Alignment Scores:
Pred. No.: 1.92e-31 Length: 1773
Score: 326.00 Matches: 59
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
Gaps: 0

SEQ5 (1-59) x AX384610 (1-1773)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 970 CAGAATGAGGACTTGGCCCGAGCCAGTCCCTTGGACAGCAGCTTCTACCGCTCAGTGCTG 1029
|||||
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 1030 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGTATACCCAGCAGGC 1089
|||||
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArg 59
|||||
Db 1090 TTCTTCTGTCAGACCTGCCCGGGGCTGGGGGATGTTGTCACACAGCAGCCGC 1146
|||||

RESULT 4
AX384608
LOCUS AX384608 1806 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 4 from Patent WO0214503.
ACCESSION AX384608
VERSION AX384608.1 GI:19577808
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .1767
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 381 a 521 c 529 g 336 t
ORIGIN

```





```

DB: 6 Gaps: 0

SEQ5 (1-59) x AR034479 (1-3768)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAATGAGACTTGGGGCCAGCCAGTCCCTTGGACAGCAGCACCTTCTACCGCTACAGTCTG 3027
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGAGTATCTGTACCCAGCAGGCG 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 3088 TTCTTCTGTCCAGACCTGCGCCGGCGCTGGGGGCGATGCTCCACACAGGACCGC 3144

RESULT 7
AX060704
LOCUS AX060704 3768 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 2 from Patent WO0100244.
ACCESSION AX060704
VERSION AX060704.1 GI:12406101
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Erickson, S. and Schwall, R.
TITLE Methods of treatment using anti-erbB antibody-maytansinoid
conjugates
JOURNAL Patent: WO 0100244-A 2 04-JAN-2001;
Genentech, Inc. (US)
FEATURES
source 1..3768
/mol_type="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN
Alignment Scores:
Pred. No.: 4,41e-31 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

SEQ5 (1-59) x AX060704 (1-3768)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAATGAGACTTGGGGCCAGCCAGTCCCTTGGACAGCAGCACCTTCTACCGCTACAGTCTG 3027
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGAGTATCTGTACCCAGCAGGCG 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 3088 TTCTTCTGTCCAGACCTGCGCCGGCGCTGGGGGCGATGCTCCACACAGGACCGC 3144

RESULT 8
AX201817
LOCUS AX201817 3768 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 1 from Patent WO0153463.
ACCESSION AX201817
VERSION AX201817.1 GI:15391666
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Cheever, M.A. and Hand-Zimmermann, S.
TITLE Compounds and methods for prevention and treatment of her-2/ neu associated malignancies
JOURNAL Patent: WO 0153463-A 1 26-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
Location/Qualifiers
source 1..3768
/mol_type="Homo sapiens"
/db_xref="taxon:9606"
CDS 1..3768
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC60126.1"
/db_xref="GI:15391667"
/translation="MELAALCRWGLLLALLPPGAASQTQVCTGTDKMLRLRPLASPETHLD
MLRHLYQGVVQVGNLELYPTNASLFDIQEYQGVYLVIAHQVQVPLQRLRIV
RGTLQFEDNVALVDNGDPLNNTTPTVGASPGGLRELQRLSLTELKGLVLIQRLRPQ
LCYQDTILWKDIFHNKNOIALTLIDTNRSRACHPCSPCMKSGRCESSEDCQSLRRT
VCAGGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLHNHSGICELCPALVTYNT
DIFESMPNPEGRYTFGASCTACPNYLSTDVGSCTLCVPLHNVEVTAEDGTRCEKC
SKPCARVCYGLGMEHLREVRVTSANIQEFAGCKIFGLSLAFELPESFDGPASNTAPL
QPEQLQVFTLEETITGLYISAMPDLSLDLVFNQLQVIRGLRHNGAYSLSLQGLGI
SWGLRSRLRELGLALIHNTHLCFVHTVPWDQLFNPHQLALLTANRPEDECVEG
LACHOLCARGHCWGPGTQCVNCSQFLRGQECVECRVLQGLPREVYNARHCLPCHPE
COPNGSVTCFGEADQCVACAHYDPPFCVACRPSGVKPDLSYMPWKFPPDEEGACQ
PGFINCHSCVDLDDKGPACQORASPLTSISANVVGILLVVLGVVGLLJRRROOKI
RKYTMRLLOETELVEPLTPSGAMPNQAMRILKTELKRVXVLGSGAGRTYKGIWI
PDGENVKIPVAIKVIRENTSRKANKEILDEAYVMAGVSPYSVRLIGLICLTSTVQLVT
QLMYGCLLDHVRENRLGSDLLNMCQIAKMSYLEDVLRHLDLAARNVLYKSP
NHVKITDFGLARLLDIDEYHADGKVPKWMALESILRR3FTHQSDVMXGYVTWE
LMTFGAKPYDGIAPAREIPDLLEKGERLPOPICTIDVYIMYKCMWIDECRPRREL
VSEFSRMAPDQRFVYVIONEDLGASPLDSTFYSLIEDDDMDGLVDAEYLVPOQGF
FCFDPAPGAGGAWHHRHSRSTSSRSGGDLTLGLEPSEEARPSPLASGASGSDVDG
DLGMAAGLQSLTHDPSLQRYSEDVPLPSETDGYVAPLTCSPQPEYVNPQDVR
POPPSPREGPLPAARPATLRLPKTLSPKNGVVKDVFAGGAVENPEYLTLPQGAA
POPHPPPAFSPAFDNLVYWDQDPPPERGAPPSTFKGTPTAENPEYLGDPV"
BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN
Alignment Scores:
Pred. No.: 4,41e-31 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

SEQ5 (1-59) x AX201817 (1-3768)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAATGAGACTTGGGGCCAGCCAGTCCCTTGGACAGCAGCACCTTCTACCGCTACAGTCTG 3027
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGAGTATCTGTACCCAGCAGGCG 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
Db 3088 TTCTTCTGTCCAGACCTGCGCCGGCGCTGGGGGCGATGCTCCACACAGGACCGC 3144

RESULT 9
AX380923
LOCUS AX380923 3768 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 9 from Patent WO0212341.
ACCESSION AX380923
VERSION AX380923.1 GI:19575767
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS Cheever, M.A. and Gheysen, D.  
TITLE Her-2/neu fusion proteins  
JOURNAL Patent: WO 0212341-A 9 14-FEB-2002;  
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)  
FEATURES  
Location/Qualifiers

1. .3768  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

CDS  
1. .3768  
/note="human Her-2/neu protein"

/codon\_start=1  
/protein\_id="CAD28392.1"  
/db\_xref="GI:19575768"

/translation="MELAALCRWGLLALLPPGAASQVCTGTMKLRLLPASPETHLD  
MLRHYGCVQVQGNLELYPTNASLSFLQDIEQVGYVLIHNRQVPLQRLRV  
RGTLQFEDNYALVDNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPQ  
LCYQDTILWKDIFHKNNQALTLDTNRSRACHPCSPMCKGSRGWGSESDCSLTPT  
VCAGCARCKGPLETDCHEQCAAGCTGPKHSDCLACILHFNHSGICELHCPALVYNT  
DTFESMPNPEGRTYFGASCVTACPNYLSLTDVGSCTLVCLPHNQEVTAEDGTQCEKC  
SKPCARVCYIGLGMHLEVRVAVTSANIQEFAGCKIFGSLAFIPESFDGDPASNTAPL  
QPEQLQVFETLEEITGYLISAWPDSLPLDSVFQNLQVIRGRILHNGAYSLSLQGLGI  
SWLGLRSLRELGLALIHNTHLFCVHTVPMDFRNPHQALLHTANRPEDECVGEG  
LACHOLCARGHCWGPPTQCVCNCSQFLRGQCEVCEVRVLOGLPREYNARHCLPCHPE  
COPONGVTCFGEADOCVACAHYKDPFCVACRCPGVKPDLSYMPWKPPDEGACQ  
PCPINCHSCVDLDDKGCPCAEQASPLTSIISAVGILLVVLGVGFGILKRRQKI  
RKYTMRLRLQETELVEPTFGSAMPNQAMRIKLETBLRKVKVLGSGAFGVYKGIWI  
PDGNNRIPVAIKVIRENTSPKANKEILDYAVNAGVSPYVSLGLGICLTSTVLQIT  
OLMPYGLLDHVRNRLGASQDILNMCQIAGMSYLEDVRLVHRDLAARNVLKSP  
NHVKTDFGLARLDIDETEHADGGKVPKIMWALESILRRRFTHQSDVNSYGVYVE  
LMTGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMLMVKCMWIDSECRPFREL  
VSESRWARDPQRVYVQNEGLGASPLDSTFYRSLLDDMDGLDAEYLVPQQGF  
FCPDPAAGAGSMVHRRSSSTRSGGDLTLGLEPSEEEAPRSLAPSEAGSGSDVFG  
DLGMAAGLQSLPTHPDPLQRYSEDPTVPLSETDGYVAPLTCSPQEVYNQPDYR  
POPPSPREGPLPAARPAAGATLERPKTLSPGKNGVYKDVFAFGGAVENPEYILTPQGGAA  
PQPHPPAFSPAFDNLYYWDQDPPERGAPSTFKGTPTAENPEYLGLDVVP"

misc\_feature

1. .1959  
/note="extracellular domain (ECD) of human Her-2/neu"

misc\_feature

205. .3765  
/note="extracellular domain (ICD) of human Her-2/neu"

misc\_feature

2968. .3765  
/note="phosphorylation domain (PD) of human Her-2/neu"

misc\_feature

2968. .3144  
/note="preferred portion of the phosphorylation domain  
(delta PD) of human Her-2/neu"

BASE COUNT 759 a 1171 c 1119 g 719 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4.41e-31 Length: 3768  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

SEQ5 (1-59) x AX380923 (1-3768)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20

Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027

Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyLeuValProGlnGlnGly 40

Db 3028 GAGACGATGACATGGGGGACCTGTGGTGTCTGAGGAGTATCTGGTACCCAGCAGGCG 3087

Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59

Db 3088 TTCTTCGTCCAGACCTGCCCCGGGCGCTGGGGGCGATGGTCCACCAGCAGCACCGC 3144

RESULT 10  
AX384604  
LOCUS AX384604 3768 bp DNA linear PAT 19-MAR-2002  
DEFINITION Sequence 1 from Patent WO0214503.  
ACCESSION AX384604  
VERSION AX384604.1 GI:19577806  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,  
McNeill, P.D. and Vedvick, T.S.

JOURNAL her-2/neu-associated malignancies  
Patent: WO 0214503-A 1 21-FEB-2002;  
CORIXA CORPORATION (US)

FEATURES  
Location/Qualifiers

1. .3768  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

CDS  
1. .3768  
/note="unnamed protein product"

/codon\_start=1  
/protein\_id="CAD28577.1"  
/db\_xref="GI:19577807"

/translation="MELAALCRWGLLALLPPGAASQVCTGTMKLRLLPASPETHLD  
MLRHYGCVQVQGNLELYPTNASLSFLQDIEQVGYVLIHNRQVPLQRLRV  
RGTLQFEDNYALVDNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPQ  
LCYQDTILWKDIFHKNNQALTLDTNRSRACHPCSPMCKGSRGWGSESDCSLTPT  
VCAGCARCKGPLETDCHEQCAAGCTGPKHSDCLACILHFNHSGICELHCPALVYNT  
DTFESMPNPEGRTYFGASCVTACPNYLSLTDVGSCTLVCLPHNQEVTAEDGTQCEKC  
SKPCARVCYIGLGMHLEVRVAVTSANIQEFAGCKIFGSLAFIPESFDGDPASNTAPL  
QPEQLQVFETLEEITGYLISAWPDSLPLDSVFQNLQVIRGRILHNGAYSLSLQGLGI  
SWLGLRSLRELGLALIHNTHLFCVHTVPMDFRNPHQALLHTANRPEDECVGEG  
LACHOLCARGHCWGPPTQCVCNCSQFLRGQCEVCEVRVLOGLPREYNARHCLPCHPE  
COPONGVTCFGEADOCVACAHYKDPFCVACRCPGVKPDLSYMPWKPPDEGACQ  
PCPINCHSCVDLDDKGCPCAEQASPLTSIISAVGILLVVLGVGFGILKRRQKI  
RKYTMRLRLQETELVEPTFGSAMPNQAMRIKLETBLRKVKVLGSGAFGVYKGIWI  
PDGNNRIPVAIKVIRENTSPKANKEILDYAVNAGVSPYVSLGLGICLTSTVLQIT  
OLMPYGLLDHVRNRLGASQDILNMCQIAGMSYLEDVRLVHRDLAARNVLKSP  
NHVKTDFGLARLDIDETEHADGGKVPKIMWALESILRRRFTHQSDVNSYGVYVE  
LMTGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMLMVKCMWIDSECRPFREL  
VSESRWARDPQRVYVQNEGLGASPLDSTFYRSLLDDMDGLDAEYLVPQQGF  
FCPDPAAGAGSMVHRRSSSTRSGGDLTLGLEPSEEEAPRSLAPSEAGSGSDVFG  
DLGMAAGLQSLPTHPDPLQRYSEDPTVPLSETDGYVAPLTCSPQEVYNQPDYR  
POPPSPREGPLPAARPAAGATLERPKTLSPGKNGVYKDVFAFGGAVENPEYILTPQGGAA  
PQPHPPAFSPAFDNLYYWDQDPPERGAPSTFKGTPTAENPEYLGLDVVP"

BASE COUNT 759 a 1171 c 1119 g 719 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4.41e-31 Length: 3768  
Score: 326.00 Matches: 59  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

SEQ5 (1-59) x AX384604 (1-3768)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20

Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027

Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyLeuValProGlnGlnGly 40

Db 3028 GAGACGATGACATGGGGGACCTGTGGTGTCTGAGGAGTATCTGGTACCCAGCAGGCG 3087

Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59

|||||

```

Db          3088 TTCTTCTGTCAGACCCCTGCCCGGGCGCTGGGGCGATGTCACACAGGCACCGC 3144
RESULT 11
AX465456
LOCUS      AX465456                      3768 bp    DNA        linear    PAT 16-JUL-2002
DEFINITION Sequence 1 from Patent WO0213847.
ACCESSION  AX465456
VERSION     AX465456.1 GI:21899819
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Gaiger, A., Cheever, M.A. and Hand-Zimmermann, S.
TITLE       Methods for diagnosis and therapy of hematological and
            virus-associated malignancies
JOURNAL     Patent: WO 0213847-A 1 21-FEB-2002;
            CORIXA CORPORATION (US)
FEATURES
            Location/Qualifiers
            source
              1..3768
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
            CDS
              1..3768
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAD42363.1"
                /db_xref="GI:21899820"
                /translation="MELAAALRWGLLLALLPPGAASQTVCVTGDMKLRLPASPETHLD
                MURHLYGQCVVQGNLELTLYPTNASLFDQIEQVGYVLIAHNQVQVPLQRLIRV
                RTQLFEDNYALVLDNGDPLNNTTPTVTGASPGRELQRLSLTELKGGVLIQRNPQ
                LCYQDTILWKDILFHKNNQALTLIDTNRSRACHPCSPMCKSGRCSESSDCQSLTRT
                VCAGGCKGKPLPTDCCHEQCAAGCTGPKHSDCLACLHNHSGICELHCPALVYNT
                DTFESMPNPEGRTFGASCVTACPYNLSTDVSGCTVLCPLNQEVTAEDGQRCCK
                SPKARCYGLGMEHLREVRATVSANIQEFAGCKIFGSLAFLPESFDGDPASNTAPL
                QPQLQVFTLEEITGYLISAMPDSLPLSVFQNLQVIRGRILHNGAYSLTIQGLGI
                SWGLRSRLHSGSLALIHNTHLCFVHTVPMDQFRNPQHALLHTANRPEDECVEG
                LACHOLCARGHCWPGPCTQVCNCSQFLRGQECVECRVLOGLPREVYNARHCLPCHE
                COPQNSCTFCGPEADQCVACAHYKDPFCVARGCPSEVKPDLSYMPWKFPDEEGACO
                PCPINTHSCVDLDDKGCAPAEQASPLTISIISAVGILLVVLGVVFGILLIRKQKI
                RYTMRLRLQELTELPELTPSGAMPNQAMRIKTELKRVKVLGSGAGFTVYKGIWI
                PDGENYKIPVAIKLRENTSPKANKEILDEAYVMAGVSPYVSRILGICLTSTVQLVT
                OLMPYKCLLDHVRNRLGSDLLNMCQIAKMSYLEDVRLVHDLAARNVLYKSP
                NHVKITDFGLARLIDDEYTHADGKVPDKWMALESILRRRFTHOSDWSYGVTVWE
                LMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYIMVWKWMDSECRPREEL
                VSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDDMDGLVDACEYLVPQGF
                FCPDPAPAGGMVHHRRSSSTRSGGDLTLGLEPSEEAAPRSLAPSGAGSDVPDG
                DLGMAAKGLQSLPHTDPSFLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQDVR
                POPPSPREGLPAARPAAGATLERKPTLSPKNGVVKDVFAGGAVENPEYLTPOGGAA
                POPHPPPAFSPAFDNLYYWDQDPPPERGAPSTFKGTPTTAENPEYLGLDVVP"

BASE COUNT      759 a 1171 c 1119 g 719 t
ORIGIN

Alignment Scores:
Pred. No.:      4,41e-31      Length:      3768
Score:          326.00      Matches:      59
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              6              Gaps:        0

SEQ5 (1-59) x AX465456 (1-3768)

QY      1 GluAspGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20
      |||||||
Db      2968 CAGAAATGAGACATTGGGCCAGCAGTCCTTGGACAGCACCTTCTACCGCTCAGTCTG 3027
      |||||||

QY      21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
      |||||||
Db      3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGAGTATCTGTTACCCACGACGCGC 3087
      |||||||

QY      41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59

```

```

Db          3088 TTCTTCTGTCAGACCCCTGCCCGGGCGCTGGGGCGATGTCACACAGGCACCGC 3144
RESULT 12
AX467229
LOCUS      AX467229                      3768 bp    DNA        linear    PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO0234287.
ACCESSION  AX467229
VERSION     AX467229.1 GI:21900511
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Beier, A.M., Gautam, A. and Mouritsen, S.R.
TITLE       Novel therapeutic vaccine formulations
JOURNAL     Patent: WO 0234287-A 3 02-MAY-2002;
            Pharmexa A/S (DK)
FEATURES
            Location/Qualifiers
            source
              1..3768
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
            CDS
              1..3768
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAD42452.1"
                /db_xref="GI:21900512"
                /translation="MELAAALRWGLLLALLPPGAASQTVCVTGDMKLRLPASPETHLD
                MURHLYGQCVVQGNLELTLYPTNASLFDQIEQVGYVLIAHNQVQVPLQRLIRV
                RTQLFEDNYALVLDNGDPLNNTTPTVTGASPGRELQRLSLTELKGGVLIQRNPQ
                LCYQDTILWKDILFHKNNQALTLIDTNRSRACHPCSPMCKSGRCSESSDCQSLTRT
                VCAGGCKGKPLPTDCCHEQCAAGCTGPKHSDCLACLHNHSGICELHCPALVYNT
                DTFESMPNPEGRTFGASCVTACPYNLSTDVSGCTVLCPLNQEVTAEDGQRCCK
                SPKARCYGLGMEHLREVRATVSANIQEFAGCKIFGSLAFLPESFDGDPASNTAPL
                QPQLQVFTLEEITGYLISAMPDSLPLSVFQNLQVIRGRILHNGAYSLTIQGLGI
                SWGLRSRLHSGSLALIHNTHLCFVHTVPMDQFRNPQHALLHTANRPEDECVEG
                LACHOLCARGHCWPGPCTQVCNCSQFLRGQECVECRVLOGLPREVYNARHCLPCHE
                COPQNSCTFCGPEADQCVACAHYKDPFCVARGCPSEVKPDLSYMPWKFPDEEGACO
                PCPINTHSCVDLDDKGCAPAEQASPLTISIISAVGILLVVLGVVFGILLIRKQKI
                RYTMRLRLQELTELPELTPSGAMPNQAMRIKTELKRVKVLGSGAGFTVYKGIWI
                PDGENYKIPVAIKLRENTSPKANKEILDEAYVMAGVSPYVSRILGICLTSTVQLVT
                OLMPYKCLLDHVRNRLGSDLLNMCQIAKMSYLEDVRLVHDLAARNVLYKSP
                NHVKITDFGLARLIDDEYTHADGKVPDKWMALESILRRRFTHOSDWSYGVTVWE
                LMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYIMVWKWMDSECRPREEL
                VSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDDMDGLVDACEYLVPQGF
                FCPDPAPAGGMVHHRRSSSTRSGGDLTLGLEPSEEAAPRSLAPSGAGSDVPDG
                DLGMAAKGLQSLPHTDPSFLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQDVR
                POPPSPREGLPAARPAAGATLERKPTLSPKNGVVKDVFAGGAVENPEYLTPOGGAA
                POPHPPPAFSPAFDNLYYWDQDPPPERGAPSTFKGTPTTAENPEYLGLDVVP"

BASE COUNT      758 a 1170 c 1121 g 719 t
ORIGIN

Alignment Scores:
Pred. No.:      4,41e-31      Length:      3768
Score:          326.00      Matches:      59
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              6              Gaps:        0

SEQ5 (1-59) x AX467229 (1-3768)

QY      1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20
      |||||||
Db      2968 CAGAAATGAGACATTGGGCCAGCAGTCCTTGGACAGCACCTTCTACCGCTCAGTCTG 3027
      |||||||

QY      21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
      |||||||
Db      3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGAGTATCTGTTACCCACGACGCGC 3087
      |||||||

QY      41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59

```

```
Db 3088 TTCCTCTGTCCAGACCTTGCCCGGCGCTGGGGGCATGGTCCACACGACCGC 3144
|||||
RESULT 13
AX481438 3768 bp DNA linear PAT 16-AUG-2002
LOCUS AX481438
DEFINITION Sequence 52 from Patent WO20055693.
ACCESSION AX481438
VERSION AX481438.1 GI:22316352
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Kreutzer, R., Limmer, S., Rost, S. and Hadwiger, P.
TITLE Method for inhibiting the expression of a target gene
JOURNAL Patent: WO 02055693-A 52 18-JUL-2002;
Ribopharma AG (DE)
FEATURES
source
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN
Alignment Scores:
Pred. No.: 4,41e-31 Length: 3768
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ5 (1-59) x AX481438 (1-3768)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 2968 CAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 3028 GAGGACGATGACATGGGGACCTGGTGGTGTGAGGAGTATCTGTACCCAGCAGGCG 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
|||||
Db 3088 TTCCTCTGTCCAGACCTTGCCCGGCGCTGGGGGCATGGTCCACACGACCGC 3144
|||||
RESULT 14
AR080259
LOCUS AR080259 4473 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 1 from patent US 5968748.
ACCESSION AR080259
VERSION AR080259.1 GI:10006994
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4473)
AUTHORS Bennett, C. Frank., Lipton, A. and Witters, L.M.
TITLE Antisense oligonucleotide modulation of human HER-2 expression
JOURNAL Patent: US 5968748-A 1 19-OCT-1999;
FEATURES
source
1..4473
/organism="unknown"
BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN
Alignment Scores:
Pred. No.: 5,32e-31 Length: 4473
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ5 (1-59) x AR080259 (1-4473)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 3142 CAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 3202 GAGGACGATGACATGGGGACCTGGTGGTGTGAGGAGTATCTGTACCCAGCAGGCG 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
|||||
Db 3262 TTCCTCTGTCCAGACCTTGCCCGGCGCTGGGGGCATGGTCCACACGACCGC 3318
|||||
Search completed: October 16, 2003, 11:03:09
Job time : 855.833 secs
```

```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ5 (1-59) x AR080259 (1-4473)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 3142 CAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 3202 GAGGACGATGACATGGGGACCTGGTGGTGTGAGGAGTATCTGTACCCAGCAGGCG 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
|||||
Db 3262 TTCCTCTGTCCAGACCTTGCCCGGCGCTGGGGGCATGGTCCACACGACCGC 3318
|||||
RESULT 15
AR167390
LOCUS AR167390 4473 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 26 from patent US 6287569.
ACCESSION AR167390
VERSION AR167390.1 GI:17903168
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4473)
AUTHORS Kipps, T.J. and Wu, Y.
TITLE Vaccines with enhanced intracellular processing
JOURNAL Patent: US 6287569-A 26 11-SEP-2001;
FEATURES
source
1..4473
/organism="unknown"
BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN
Alignment Scores:
Pred. No.: 5,32e-31 Length: 4473
Score: 326.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ5 (1-59) x AR167390 (1-4473)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
|||||
Db 3142 CAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
|||||
Db 3202 GAGGACGATGACATGGGGACCTGGTGGTGTGAGGAGTATCTGTACCCAGCAGGCG 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg 59
|||||
Db 3262 TTCCTCTGTCCAGACCTTGCCCGGCGCTGGGGGCATGGTCCACACGACCGC 3318
|||||
Search completed: October 16, 2003, 11:03:09
Job time : 855.833 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 2081.84 Seconds  
(without alignments)  
3105.423 Million cell updates/sec

Title: SEQ4  
Perfect score: 1450  
Sequence: 1 QNEDLGPASPLDSTFYRSL.....TFKGTPTAENPEYLGLDVPV 266

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool.p/HOLLERAN480/runat\_15102003\_131913\_20548/app\_query.fasta\_1.4685  
-DB=EST\_QFMT-fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -TRANS-human40.cdi -LIST=45  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -FRAMES=human40.cdi -ALIGN=15 -MODE=LOCAL  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=HOLLERAN480 -CGCN\_1\_12645 -runat\_15102003\_131913\_20548 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estinu:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_fod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gssl:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1426	98.3	885	14	CA455074	CA455074 AGENCOURT
2	1423	98.1	4715	11	AF318349	AF318349 Homo sapi
3	1399.5	96.5	893	14	CA455141	CA455141 AGENCOURT
4	1354	93.4	894	14	CA454131	CA454131 AGENCOURT
5	1340	92.4	1004	14	CD515356	CD515356 AGENCOURT
6	1292	89.1	1129	12	BM802792	BM802792 AGENCOURT
7	1279	88.2	871	14	CA488274	CA488274 AGENCOURT
8	1205	83.1	1988	11	BC023725	BC023725 Mus muscu
9	1205	83.1	3110	11	AK031542	AK031542 Mus muscu
10	1205	83.1	3372	11	BC046553	BC046553 Mus muscu
11	1205	83.1	4323	11	AK031099	AK031099 Mus muscu
12	1205	83.1	4463	11	AK083669	AK083669 Mus muscu
13	1165.5	80.4	852	14	CD516283	CD516283 AGENCOURT
14	1136.5	78.4	932	14	CA487981	CA487981 AGENCOURT
15	1127	77.7	685	9	AU123871	AU123871 AU123871
16	1092	75.3	653	12	BM721340	BM721340 UI-E-E01-
17	1084	74.8	902	14	CA488688	CA488688 AGENCOURT
18	1070	73.8	791	12	BI154872	BI154872 602902857
19	1035	71.4	887	14	CA980253	CA980253 AGENCOURT
20	1032	71.2	943	13	BQ958632	BQ958632 AGENCOURT
21	966	66.6	919	13	BQ958632	BQ958632 AGENCOURT
22	954	65.8	537	12	BM793744	BM793744 K-EST0074
23	936	64.6	813	12	BI557797	BI557797 603236977
24	933	64.3	899	14	CD558386	CD558386 AGENCOURT
25	925	63.8	855	13	BU594980	BU594980 AGENCOURT
26	915	63.1	717	14	CB598701	CB598701 AGENCOURT
27	899.5	62.0	916	13	BU845369	BU845369 AGENCOURT
28	898	61.9	906	14	CA454570	CA454570 AGENCOURT
29	897	61.9	547	12	BM787824	BM787824 K-EST0066
30	896	61.8	491	12	BM790293	BM790293 K-EST0070
31	889	61.3	998	12	BI649877	BI649877 603296516
32	880.5	60.7	787	12	BI156730	BI156730 602922150
33	880.5	60.7	844	12	BI158626	BI158626 602922395
34	880	60.3	609	9	AW701942	AW701942 ug93b02.y
35	874	60.3	621	10	BG283493	BG283493 602407782
36	861.5	59.4	849	14	CA489799	CA489799 AGENCOURT
37	860	59.3	786	12	BI155788	BI155788 602904360
38	858	59.2	887	13	BQ717097	BQ717097 AGENCOURT
39	843.5	58.2	649	9	AA496412	AA496412 zv37a02.r
40	829.5	57.2	629	10	BE374596	BE374596 601225384
41	829	57.2	533	14	CB437620	CB437620 685491.MA
42	829	57.2	552	10	BG277542	BG277542 ux45d07.y
43	817	56.3	518	10	BF039952	BF039952 BP250015A
44	809	55.8	575	12	BG70469	BG70469 602838347
45	804.5	55.5	678	12	BI555157	BI555157 603236396

ALIGNMENTS

RESULT 1  
CA455074  
LOCUS  
DEFINITION AGENCOURT\_10735919 MAPCL Homo sapiens cdna clone IMAGE:6722585 5',  
885 bp mRNA linear EST 12-NOV-2002  
CA455074 mRNA sequence.  
ACCESSION CA455074  
VERSION CA455074.1 GI:24905427  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 885)  
gb\_gssl:\*

**AUTHORS** NTH-MGC <http://mgc.nci.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs@email.nih.gov](mailto:cgapbs@email.nih.gov)  
 Tissue Procurement: Kristi A. Eglund, Ira Pastan  
 cDNA Library Preparation: Invitrogen Corp  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LENL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LENL at:  
<http://image.llnl.gov>  
 Plate: L1AMI4285 row: 1 column: 17  
 High quality sequence stop: 717.

**FEATURES**  
 source  
 1..885  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6722585"  
 /cell\_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1  
 , LNCap"  
 /lab\_host="EMDH10B"  
 /clone\_lib="MAPCL"  
 /note="Vector: pCMV-SPORT6; Site\_1: EcoRV; Site\_2: Not I;  
 Subtracted with brain, liver, lung, kidney and muscle.  
 Directionally cloned. Priming method: oligo-dt. Average  
 insert size: 1800 bp. Library amplification: 26,000 fold.  
 Kristi A. Eglund, James J. Vincent, Robert Strausberg,  
 Bungkok Lee & Ira Pastan: Discovery of new breast  
 cancer genes encoding membrane and secreted proteins.  
 Manuscript submitted."

**BASE COUNT** 178 a 297 c 261 g 149 t

**ORIGIN**

**Alignment Scores:**  
 Pred. No.: 1.61e-79 Length: 885  
 Score: 1426.00 Matches: 264  
 Percent Similarity: 99.62% Conservative: 0  
 Best Local Similarity: 99.62% Mismatches: 0  
 Query Match: 98.34% Indels: 1  
 DB: 14 Gaps: 0

SEQ4 (1-266) x CA455074 (1-885)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyArgSerLeuLeu 20  
 Db 1 CAGANTGAGGACTTGGGCCAGCCAGCTCCCTTGACAGCACCTCTACCGCTCACGTGCTG 60

Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluLeuValProGlnGlnGly 40  
 Db 61 GAGGACGATCACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 120

Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60  
 Db 121 TTCTTCTGTCCAGACCTCGCCCGGGCGCTGGGGGATGCTCCACACAGCCAGCCAGC 180

Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
 Db 181 TCATCTACAGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 240

Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
 Db 241 GCCCCAGGCTCCACTGGCACCTCCGAAGGGGCTGGCTCCGATGATATTGTGATGTCAC 300

Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
 Db 301 CTGGGAATGGGGCAGCAGCAAGGGGCTGCAAGGCTCCCAACATGACCCACCCCTCTA 360

Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
 Db 361 CAGCGGTACAGTCAGGAGCCCAAGTACCCCTGCTGAGACTGATGGCTACGTTGCC 420

Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
 Db 421 CCCCTGACCTGCAGCCGCCAGCCTGAATATGTGAACAGCCAGATGTTGCGCCGCCAGCC 480

Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
 Db 481 CCTTCGCCCGAGAGGGCCCTCTGCTGCTGCGCCAGCTGCTGCTGCTGCTGCTGCTGCTG 540

Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
 Db 541 CCCAAGACTCTCTCCCCAGGGAAGATGGGGTCTCAAGACAGCTTTTGTCTTTGGGGGT 600

Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220  
 Db 601 GCGGTGGAGAACCCCGAGTACTTGCACACCCAGGAGGAGCTGCCCTTCAGCCGCCACCT 660

Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTrpAspGlnAspProProGlu 240  
 Db 661 CCTCCTGCTTTCAGCCCGAGCTTCGACAACTCTATTACTTGGACGAGCCAGCCAGCAG 720

Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
 Db 721 CGGGGGCTCCACCCAGCACCTTCAAGGGACACCTACGCGAGAACCCAGAGTACCTG 780

Qy 261 -GlyLeuAspVal 264  
 Db 781 GGCTCTGGAGGTG 793

**RESULT 2**  
 AF318349  
 LOCUS Homo sapiens pp3659 mRNA, complete cds.  
 DEFINITION AF318349  
 ACCESSION AF318349.1 GI:18027789  
 VERSION  
 KEYWORDS HTC.  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 4715)  
 AUTHORS Huang, Y., Zhou, X.M., Zhang, P.P., Jiang, H.Q., Qin, W.X., Zhao, X.T.,  
 Wan, D.F. and Gu, J.R.  
 TITLE Novel human cDNA clones with function of inhibiting cancer cell  
 growth  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 4715)  
 AUTHORS Huang, Y., Zhou, X.M., Zhang, P.P., Jiang, H.Q., Qin, W.X., Zhao, X.T.,  
 Wan, D.F. and Gu, J.R.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-NOV-2000) National Laboratory For Oncogenes & Related  
 Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai  
 200032, P. R. China  
 FEATURES  
 source  
 1..4715  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 gene  
 1..4715  
 /gene="pp3659"  
 CDS  
 2950..4188  
 /gene="pp3659"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AAL55856.1"  
 /db\_xref="GI:18027790"  
 /translation="MALESILRRFRTHQSDVWSYGVTVWELMTFGAKPYDGPARELP  
 DLLEKGRPQPDPICITDVMIMVKMIDSECRPRELVSSEFSRNPDRQFVVIQ  
 NEDLGASPLDSTFYRSLLEDMDGLVDAEYLVVQQGFPCDPAPAGGMVHRRH  
 SSSTRSGGGDLTLGLPSEEAAPSLAPSEGAGSDVFDGLGMAKGLQSLPTDHP  
 SPQRYSEDTVPPLSETDGYVAPLTCSPQEVYNQDVRPQPPSPREGPLPAARPAG  
 ATLERPKTSLSPCKNGVVKDVFAGGAVENPEYLTPOGGAALSPDLLPSAPSTSTT  
 KTRTHSGGLHPAPSKGHLNRQSTWVWTCQCEPQGVRRSPDVSSSGSREGLTSGI  
 KRWEGPPTTSRGTCARNLS"



BASE COUNT 969 a 1446 c 1347 g 953 t  
ORIGIN

Alignment Scores:

Pred. No.: 1, 21e-78 Length: 4715  
Score: 1423.00 Matches: 265  
Percent Similarity: 99.62% Conservativeness: 0  
Best Local Similarity: 99.62% Mismatches: 1  
Query Match: 98.14% Indels: 1  
DB: 11 Gaps: 0

SEQ4 (1-266) x AF318349 (1-4715)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
Db 3253 CAGATGAGGACTGGGGCCAGCCAGTCCCTTGACAGCAGCTTCTACCGCTCACTGGTG 3312  
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40  
Db 3313 GAGGACGATGATGGGGGACCTGGTGATCTGAGGAGTATCTGGTACCCAGCAGGGC 3372  
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgSer 60  
Db 3373 TTCCTCTGTCAGACCTGCCCCGGGGCGCTGGGGGCGATGGTCCACAGCAGCCGCGAGC 3432  
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
Db 3433 TCATCTACGAGAGTGGGGTGGGACCTGACATGAGGCTGGAGCCCTCTGAAGAGGAG 3492  
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
Db 3493 GCCCCAGGTCTCCAGTGGCACCTCCGAAGGGGCTGGCTCCGATGTATTGTGTGTCAC 3552  
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
Db 3553 CTGGGAATGGGGGAGCCAGAGGGCTGCAAGAGCTCCACACATGACCCAGCCCTCTA 3612  
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
Db 3613 CAGGGTACATGAGGACCCACAGTACCTCCCTCTGAGACTGATGGCTAGCTGCC 3672  
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
Db 3673 CCCCTGACCTGCAGCCCCAGCCCTGAATATGTGAACACAGCCAGATGTTTCGCCCCAGGCC 3732  
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
Db 3733 CCTTGGCCCCGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGTCACACTCTGGAAGG 3792  
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
Db 3793 CCCAAGACTCTCTCCAGGAGGAGAGTGGGGTCTGCAAGAGCGTTTTTGGCCCTTGGGGGT 3852  
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220  
Db 3853 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGTGC-CCTCAGCCCCACCT 3911  
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240  
Db 3912 CCTCTGCTTTCAGCCCGACCTTGGACAACTCTATTACTGGGACGAGGCCACCCAGAG 3971  
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
Db 3972 CGGGGGGCTCCACCCAGACCTTCAAGGGGACCTTACGGGAGAGAACCCAGAGTACCTG 4031  
Qy 261 GlyLeuAspValProVal 266  
Db 4032 GGTCTGGACGTGCCAGT 4049

RESULT 3

CA455141 893 bp mRNA linear EST 12-NOV-2002  
LOCUS CA455141  
DEFINITION AGENCOURT\_10735980 MAPCL Homo sapiens cdna clone IMAGE:6722663 5',

mRNA sequence.

ACCESSION CA455141  
VERSION CA455141.1 GI:24905561  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM

REFERENCE 1 (bases 1 to 893)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: Kristi A. Eglund, Ira Pastan  
CDNA Library Preparation: Invitrogen Corp  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: LLAM14285 row: 1 column: 23  
High quality sequence stop: 682.  
Location/Qualifiers  
1..893

FEATURES  
source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6722663"  
/cell\_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1", LNCap  
/lab\_host="EMDH108"  
/clone\_lib="MAPCL"  
/note="Vector: PCMV-SPORT6; Site\_1: EcoRV; Site\_2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dt. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg, Bungkook Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

BASE COUNT 175 a 302 c 261 g 155 t  
ORIGIN

Alignment Scores:

Pred. No.: 7, 21e-78 Length: 893  
Score: 1399.50 Matches: 261  
Percent Similarity: 98.13% Conservativeness: 1  
Best Local Similarity: 97.75% Mismatches: 3  
Query Match: 96.52% Indels: 2  
DB: 14 Gaps: 1

SEQ4 (1-266) x CA455141 (1-893)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
Db 67 CAGATGAGGACTGGGGCCAGCCAGTCCCTTGACAGCAGCTTCTACCGCTCACTGGTG 126  
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40  
Db 127 GAGGACGATGATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 186  
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgSer 60  
Db 187 TTCCTCTGTCAGACCTGCCCCGGGGCGCTGGGGGCGATGGTCCACAGCAGCCGCGAGC 246  
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
Db 247 TCATCTACGAGAGTGGGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 306  
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
Db 307 GCCCCAGGTCTCCACTGGCACCCCTCCGAAGGGGCTGCTCCGATGTATTGTGATGTCAC 366

```

101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
111 |
121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
131 |
141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
151 |
161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
171 |
181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
191 |
201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
211 |
221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
231 |
241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLe 260
251 |
260 uGlyLeu---AspValPro 265
271 |
281 TGGGGTCTGGGACGTGCC 865

RESULT 4
CA454131
LOCUS AGENCOURT_10738550 MAPcL Homo sapiens cDNA clone IMAGE:6718792 5',
DEFINITION mRNA sequence.
ACCESSION CA454131
VERSION CA454131.1 GI:24903569
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 894)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
cDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL14275 row: k column: 16
High quality sequence stop: 636.
FEATURES
source
1..894
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6718792"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HMEL
, LNCaP"
/lab_host="EMDH10B"

```

```

/clone_lib="MAPcL"
/notes="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."
BASE COUNT 180 a 303 c 260 g 151 t
ORIGIN
Alignment Scores:
Pred. No.: 4.86e-75 Length: 894
Score: 1354.00 Matches: 262
Percent Similarity: 96.69% Conservative: 1
Best Local Similarity: 96.32% Mismatches: 3
Query Match: 93.38% Indels: 6
DB: 14 Gaps: 0
SEQ4 (1-266) x CA454131 (1-894)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 24 CAGAAATGAGGACTTGGGCCCGACGTCCTTGGACAGCACCTTCTACCGCTCACTGCTG 83
QY 21 GlnAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
DB 84 GAGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 143
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
DB 144 TTCTTCTGTCCAGACCTGCCCCGGCGCTGGGGCATGCTGCCACACAGCAGCCGCGAG 203
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
DB 204 TCATCTACCCAGGAGTGGCGGTGGGACCTAGGCTGAGGCTGAGGCTCTGAAGAGGAG 263
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
DB 264 GCCCCAGGCTCTCCACTGGCACCTCCGAAGGGGCTGGCTCCGATGATTTGATGTGAC 323
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
DB 324 CTGGGAATGGGGCAGCCAAAGGGCTGCCAAAGCTCCCCACACATGATGCCAGCCCTCTA 383
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
DB 384 CAGCGGTACAGTACGAGGACCCACAGTACCCCTCTGAGACTGATGGCTACGTTGCC 443
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
DB 444 CCCTGACCTGCAGCCCCCAGCCCTGATATGTAACAGCAGCATGTTCCGGCCCCAGCCC 503
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
DB 504 CCTTCGCCCGCAGAGGGCCCTGCTGCTGCTGCCGACCTGCTGGTGCACCTCTGGAAAGG 563
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
DB 564 CCCAAGACTCTCTCCCCAGGGAAGAATGGGGTCTGCTCAAGACACGTTTTTGGCTTTGGGGGT 623
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
DB 624 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCCCTCAGCCCCACCC 683
QY 220 OProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrp-AspGlnAspProPro 240
DB 684 TCCTCTGCTTCCAGCCCGACCTTCGACAACTCTATTACTGGGACCCAGGACCCCGAGG 743
QY 240 LuArgGlyAla--ProProSerThrPheLysGlyThrProThrAlaGluAsnProGlu-T 259
DB 744 AGCGGGGGGGCTCCCGACCCAGCACCTTTTCAAGAGGACACTACGCGCAGAGAACCAGAGT 803

```

[illegible]

```

FEATURES
source
Location/Qualifiers
1..1129
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5575087"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 243 a 367 c 325 g 194 t
ORIGIN
Alignment Scores:
Pred. No.: 4,33e-71 Length: 1129
Score: 1292.00 Matches: 239
Percent Similarity: 99.17% Conservative: 0
Best Local Similarity: 99.17% Mismatches: 1
Query Match: 89.10% Indels: 1
DB: 12 Gaps: 0

SEQ4 (1-266) x BM802792 (1-1129)
QY 27 AspLeuValAspAlaGluGluTyrLeuValProGlnGlnGlyPheCysProAspPro 46
DB 3 GACCTGGTGATGCTGAGGATGTCGTACCCACGACGAGGCTTCTGTCTCCAGACCT 62
QY 47 AlaProGlyAlaGlyGlyMetValHisArgHisArgSerSerThrArgSergly 66
DB 63 GCCCGGGGGCTGGGGGCTGTCACACGACGACGACGACGACGACGACGACGACG 122
QY 67 GlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeu 86
DB 123 GGTGGGGACCTGACATAGGCTGGAGCCCTCTGAAGAGAGAGAGCCGCCAGGTCTCCACTG 182
QY 87 AlaProSerGluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAla 106
DB 183 GCACCTCCGAGGGGCTGGCTCCGATGATTTGATGGTGACCTGGGATGGGGGAGCC 242
QY 107 LysGlyLeuGlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAsp 126
DB 243 AAGGGGCTGCAAGGCTCCGCCACATGACCCAGCCCTCTACAGCGGTACAGTGAGGAC 302
QY 127 ProThrValProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerPro 146
DB 303 CCCACAGTACCCCTGCGCTGAGACTGATGGCTAGCTGCCCCCTGACCTGCAGCCCC 362
QY 147 GlnProGluTyrValAsnGlnProAspValArgProGlnProProSerProArgGluGly 166
DB 363 CAGCGTGAATATGTAACCCAGCAGATGTCGCCCCAGCCCTTCGCCCCGAGAGGGC 422
QY 167 ProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerPro 186
DB 423 CCTCTGCCTGCTCCCGACCTGCTGGTGCACCTCTGGAAGGGCCCAAGACTCTCTCCCA 482
QY 187 GlyLysAsnGlyValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGlu 206
DB 483 GGAAGAATGGGCTGCTCAAGACGTTTTCCTTTGGGGGTGCCGTGGAGAACCCCGAG 542
QY 207 TyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerPro 226
DB 543 TACTTGACACCCAGGAGAGCTGCCCTCAGCCCCACCTCTCTCTGCTTCAGCCCA 602
QY 227 AlaPheAsnLeuTyrTyrTrpAspGlnAspProGluArgGlyAlaProProSer 246
DB 603 GCCTTCGACACCTCTATTACTGGGACGAGGACCCACGAGGGGGGCTCCACCCAGC 662
QY 247 ThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu-GlyLeuAspValProVa 266

```

---

```

Db 663 ACCTTCAAGGGGACACCTAGCGCAGAGAACCCAGAGTACCTGGGCTTGACGTGCCAGT 722
QY 266 1 266
Db 723 G 723
RESULT 7
CA488274 871 bp mRNA linear EST 14-NOV-2002
LOCUS AGENCOURT_10808010 MAPCL Homo sapiens cDNA clone IMAGE:6720019 5',
DEFINITION mRNA sequence.
ACCESSION CA488274
VERSION CA488274.1 GI:24949701
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 871)
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-femail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14278 row: n column: 19
High quality sequence stop: 693.
FEATURES
source
Location/Qualifiers
1..871
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6720019"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HMEL1, LNCaP"
/lab_host="EMDH10B"
/clone_lib="MAPCL"
/notes="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."
BASE COUNT 177 a 296 c 252 g 144 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 2.18e-70 Length: 871
Score: 1279.00 Matches: 236
Percent Similarity: 99.16% Conservative: 0
Best Local Similarity: 99.16% Mismatches: 1
Query Match: 88.21% Indels: 1
DB: 14 Gaps: 0

SEQ4 (1-266) x CA488274 (1-871)
QY 30 AspAlaGluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGly 49
DB 3 GATGCTGAGAGATCTGGTACCCAGCAGGGCTTCTGTCTCCAGACCTGCCCGGGC 62
QY 50 AlaGlyGlyMetValHisArgHisArgSerSerThrArgSerGlyGlyAsp 69
DB 63 GCTGGGGCATGCTCCACCACGACGACGACGACGACGACGACGACGACGACGACG 122
QY 70 LeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSer 89

```

```

|||||
123 CTGACACTAGGGCTGGAGCCCTCTCAAGAGAGGAGCCCGCCAGGCTCCACCTGCACCCCTCC 182
QY 90 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyGlyLeu 109
Db 183 GAAGGGGTGGCTCCGATGATTTGATGCTGACCTGGGAATGGGGCCAGCCAAAGGGGCTG 242
QY 110 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 129
Db 243 CAAAGCTCCACACATGACCCAGCCCTCTACAGCGGTGACAGTGAGAGACCCACAGTA 302
QY 130 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGlu 149
Db 303 CCCCTGCCCTCTGAGACTGATGGCTACGTTGCCCTGACCTGACCTGCAGCCCGAGCTGAA 362
QY 150 TyrValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuPro 169
Db 363 TATGTGAACACAGCCAGATGTTGGGCCCGAGCCCTGCGCCCGAGAGGGCCCTCGCT 422
QY 170 AlaAlaArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsn 189
Db 423 GCTGCCACACCTGCTGGTGGCCACTCTGGAAGGCCACAGACTCTCTCCCGAGGGAAGAT 482
QY 190 GlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 209
Db 483 GGGGTGCTCAAGACGCTTTTGGCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACA 542
QY 210 ProGlnGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 229
Db 543 CCCAGGAGGAGCTGCCCTCAGCCACCCCTCTCTGCTTCCAGCCAGCTTCGAC 602
QY 230 AsnLeuTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyr 249
Db 603 AACCTCTATTACTGGACAGGAGCCACAGAGCGGGGGCTCCACCCAGCACCTTCAA 662
QY 250 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 266
Db 663 GGGACACCTACGGCAGAGAACCCAGACTACCTGGNGTCTGGAGCTGCCAGTG 714

RESULT 8
BC023725 1888 bp mRNA linear HTC 04-MAR-2003
LOCUS
DEFINITION
Mus musculus, Similar to Avian erythroblastosis viral (v-erb-B2)
oncogene homologue 2 (neuro/glioblastoma derived oncogene homolog),
clone IMAGE:5347334, mRNA.
ACCESSION
BC023725
VERSION
BC023725.1 GI:23959125.
KEYWORDS
HTC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1888)
Strausberg,R.
Direct Submission
Submitted (05-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
REMARK
COMMENT
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 55 Row: 1 Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: retained intron.
FEATURES
Location/Qualifiers
source
1..1888
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5347334"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI_CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT 444 a 580 c 532 g 432 t
ORIGIN
Alignment Scores:
Pred. No.: 1..89e-65 Length: 1888
Score: 1205.00 Matches: 223
Percent Similarity: 87.9% Conservative: 11
Best Local Similarity: 83.8% Mismatches: 32
Query Match: 83.1% Indels: 0
DB: 11 Gaps: 0
SEQ4 (1-266) x BC023725 (1-1888)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 123 CAGAACGAGGACTTAGGCCCTCCAGCCCCATGGAGACACCTTCTACCGTTCTACGTGCTG 182
QY 21 GluAspAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 183 GAGGATGATGACATGGGGAGCTGCTGATGCTGAAGAGTACTGTTACCCAGCAGGGA 242
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgSer 60
Db 243 TTCTTCTCCAGACCTTGCCTAGGTACTGGGAGCAGACAGCCAGCAGACCGCAGC 302
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGlyProSerGluGluGlu 80
Db 303 TCGTCGGCCAGAGTGGCGGTGGTGGCTGACACTGGGCTCGGAGCCCTCGGAAAGAG 362
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 363 CCCCCACATCTCCACTGGCTCCCTCCGAGGGGCTGGCTCCGATGCTGTTGATGGTAC 422
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 423 CTGGCAGTGGGGGTAAACCAAGAGACTGACAGAGCTCTCTCCACATGACCTCAGCCCTCA 482
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 483 CAGCGGTACAGTGAGGATCCCATTTACTCTGCCCCCGAGAGCTGATGGGTGCTTGTCT 542
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 543 CCCCTGGCTGCAGGCCCGCCAGCCAGTATGTGAACACAGAGAGTTCGGCCTCAGTCT 602
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 603 CCCTTGACCCAGAGGGTCTCCCGCTCCCATCCGACCTGCTGGTGGTGTCTACTCTAGAA 662

```



SSSCVARGSPGVKPDLSYMPIWKVPDEEGICQPCPINCTHSCVDLDERGCPAERASP  
 QFIATVGVLLFLIIIVVIGILIKRRRQIRKTYMRRLIQETELVPLTPSPGAVP  
 VQOMELATELRLKLVKGAGFTVYKGIWIPDGENVKIIPALKVLRNTPSPRANKE  
 ILDEAYVWAGVSPVSRLLGICLTSTVQLTQVLPYCCLLDHVREHGRGRSQDQLN  
 WCQVIAKMSILEEVRVLRDLAARNVLKSPNVHKIINDFGLARLLDIDETFEYHADGG  
 KVPITKMALESILRRRFTHQSDVMSYGVTVWELTFGAKPYDGIPIRIPDLLEKGER  
 LQPQFICVIMVIMVKWMDSECRPRFRELVSFRMARDPQRFVQIONEDLGPS  
 PMDSFYFSLLEDODDDELVAEYLVPOQGFSPDPALGTGTGAHRRHRSRSARSGG  
 GELTGLSPSEEPSPPLAPSEAGSDVFDGDLAVGTGKGLSLSPHDLPLORYSE  
 DPTLPLPETGCVYAPLACSPQPEYVNOPEVRPQSPPLPPEGPPPIRIPAGATLERPKT  
 LSPGKNGVKKDVFAPGAVENPEYLAPRAGTASOPHPSPAFSPAFDNLYYWDQNSSEQ  
 GPPSTFEFTTAENPEYLGIDVPV\*

BASE COUNT 677 a 886 c 891 g 656 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2,89e-65 Length: 3110  
 Score: 1205.00 Matches: 223  
 Percent Similarity: 87.97% Conservative: 11  
 Best Local Similarity: 83.83% Mismatches: 32  
 Query Match: 83.10% Indels: 0  
 DB: 11 Gaps: 0

SEQ4 (1-266) x AK031542 (1-3110)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 DB 1848 CAGACGAGGACTTAGGCCCTCCAGCCCCATGGACACACCTTCTACCGTCTACGTG 1907

QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40  
 DB 1908 GAGGATGATGACATGGGGAGCTGCTGATGCTGAAGAGTACCTGGTACCCAGCAGGGA 1967

QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60  
 DB 1968 TTCTTCTCCACAGACCTGCGCCTAGGTACTGGAGCACAGCCACCGACACCGCCAGC 2027

QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
 DB 2028 TCGTCGGCAGGAGTGGGGTGGTGAGTGACATGGCGCTGGAGCCCTCGGAAGAAGAG 2087

QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
 DB 2088 CCCCCAGATCTCCACTGGCTCCCTCCGAAAGGGCTGGCTCGGATGTGTTGATGCTGAC 2147

QY 101 LeuGlyMetGlyAlaAlaGlySerLeuGlnSerLeuProThrHisAspProSerProLeu 120  
 DB 2148 CTGGCAGTGGGGGTAAACCAAGGACTGCAGAGCCCTCTCCACATGACCTCAGCCCTCTA 2207

QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
 DB 2208 CAGCGGTACAGTGAAGATCCACATTAATCTCTGCCCCCGAGAGCTGATGGCTACGTGCT 2267

QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
 DB 2268 CCCCCTGGCTGCAGCCCCCAGCCGAGTATGTGAACACGACGAGGTTTCGGCCTCAGTCT 2327

QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
 DB 2328 CCCCCTGACCCAGAGGGGCTCCCGCCTCCCATCCGACCTGCTGGTGTACTCTAGAAAGA 2387

QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
 DB 2388 CCAAGACTCTCTCTCCGGGAAATGGGGTGTGAAGAGCTTTTTCCTCTGGGGGT 2447

QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220  
 DB 2448 GCTGTGGAGAACCTGAATACTTAGCACCCAGAGCAGGACTGCTCTCAGCCCCACCT 2507

QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240  
 DB 2508 TCTCTGCTTCCAGCCCGACCTTTTGACAACTCTATTACTGGGACCAAGAACTCATCGGAG 2567

QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
 DB 2568 CAGGCTCTCCACCAGTACCTTTGAAGGGACCCCTGCAGAGACCTCGAGTACCTA 2627

QY 261 GlyLeuAspValProVal 266  
 DB 2628 GGCCTGGATGTCAGCA 2645

RESULT 10  
 BC046553 3372 bp mRNA linear HTC 03-FEB-2003  
 LOCUS Mus musculus, clone IMAGE:5340777, mRNA.  
 DEFINITION BC046553  
 ACCESSION BC046553.1 GI:28196923  
 VERSION HTC.  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 3372)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-FEB-2003) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov  
 COMMENT Contact: MGC help desk  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
 Contact: amg@bcm.tmc.edu  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,  
 Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Series: IRAK Plate: 92 Row: n Column: 11  
 This clone has the following problem: no 5' EST match.

FEATURES  
 source

1..3372  
 /location=Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129,C57BL/6J,FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5340777"  
 /tissue\_type="Mammary tumor. Brca1-/fl; MMTV-Cre model. 10  
 months old, gross tissue."  
 /clone\_lib="NCI CGAP\_Mam3"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 BASE COUNT 754 a 944 c 926 g 746 t  
 ORIGIN

Alignment Scores:

Pred. No.: 3,12e-65 Length: 3372  
 Score: 1205.00 Matches: 223  
 Percent Similarity: 87.97% Conservative: 11  
 Best Local Similarity: 83.83% Mismatches: 32  
 Query Match: 83.10% Indels: 0  
 DB: 11 Gaps: 0

SEQ4 (1-266) x BC046553 (1-3372)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 DB 1505 CAGACGAGGACTTAGGCCCTCCAGCCCCATGGACACACCTTCTACCGTCTACGTG 1564







Fletcher,C., Fujita,M., Gariboldi,M., Gustinich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S., and Hayashizaki,Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

21085660

11217851

5

The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 4463)

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

#### FEATURES

source

Location/Qualifiers  
1..4463  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM,DB:D030063B12"  
/db\_xref="taxon:10090"  
/clone="D030063B12"  
/tissue\_type="whole body"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="9 days embryo"

misc\_feature

1..4463  
/note="v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian) (MGDI|MG1:95410, GB|U71126, evidence: BLASTN, 99%, match=449)";

BASE COUNT 950 a 1290 c 1270 g 953 t

ORIGIN

Alignment Scores:

Pred. No.: 4,07e-65 Length: 4463

Score: 1205.00 Matches: 223

Percent Similarity: 87.97% Conservative: 11

Best Local Similarity: 83.83% Mismatches: 32

Query Match: 83.10% Indels: 0

DB: 11 Gaps: 0

SEQ4 (1-266) x AK083669 (1-4463)

QY 1 GUAAGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
|||||  
Db 3100 CAGAACGAGGACCTAGGCCCCCTCCAGCCCCGACACCTTCTACCGTTCACTGCTG 3159  
QY 21 GUAAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40  
|||||  
Db 3160 GAGGATGATCAGATGGGGAGCTGGTGCTGAAGAGTACCTGGTACCCAGCAGGGA 3219  
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60  
|||||  
Db 3220 TTCTTCTCCACAGACCTGCTGCTAGGTACTGGGAGCAGACGCCACCCAGACACCGCAGC 3279  
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
|||||  
Db 3280 TCGTGGCCAGGAGTGGCGGTGGAGCTGACACTGGGCGCTGGAGCCCTCGGAAGAAGAG 3339  
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
|||||  
Db 3340 CCCCCAGATCTCCACTGGCTCCCTCCGAAAGGGCTGGCTCCGATGNGTTGATGGTGAC 3399  
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
|||||  
Db 3400 CTGGCAGTGGGGGTAAACCAAGGACTGCAGAGCCTCTCTCCACATGACCTCAGCCCTCTA 3459  
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
|||||  
Db 3460 CAGCGGTACAGTGAAGTCCACATACCTCTGCTGCCCGCCGAGACTGATGGCTAGTGCT 3519  
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
|||||  
Db 3520 CCCCTGGCCTGACAGCCCCAGCCGAGTATGTGAACCCAGCAGAGGTTGCGCCTCACTCT 3579  
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
|||||  
Db 3580 CCCTTGACCCAGAGGGTCTCCGCTCCCATCCGACCTGCTGGTGTACTCTATAAGA 3639  
QY 181 ProlysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
|||||  
Db 3640 CCCAAGACTCTCTCTCTGGGAAAATGGGGTTGTCAAAGACGTTTTGCTTTGGGGGT 3699  
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220  
|||||  
Db 3700 GCTGTGAGAACCTGAATATTCTAGCACCCAGCAGCAGGCTGCTCTCAGCCCCCACCCT 3759  
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240  
|||||  
Db 3760 TCTCCTGCCTTCAGCCAGCCTTTGACAACCTCTATTACTGGGACCAAGACTCATCGGAG 3819  
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
|||||  
Db 3820 CAGGTCCTCCACCAAGTACTTTGAAGGGGACCCCACTGCAGAGAACCTTGAGTACCTA 3879  
QY 261 GlyLeuAspValProVal 266  
|||||  
Db 3880 GGCCTGATGTCAGTA 3897

RESULT 13  
CD516283  
LOCUS CD516283  
DEFINITION AGENCOURT\_14364504 NIH\_MGC\_181 Homo sapiens CDNA clone  
IMAGE:30408904 5', mRNA sequence.  
ACCESSION CD516283  
VERSION CD516283.1 GI:31448001  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 852)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE  
JOURNAL  
COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics / NIH  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein  
cDNA Library Preparation: Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
plate: NDAM492 row: j column: 17  
High quality sequence stop: 662.  
Location/Qualifiers  
1. 852  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30408904"  
/tissue\_type="White Matter"  
/dev\_stage="Unknown"  
/lab\_host="DH10B-Ton A ( T1 and T5 phage resistances)"  
/clone\_lib="NIH\_MGC\_181"  
/note="Vector: pCMV-SPORT6.1; Site\_1: NotI; Site\_2: EcoRV  
(destroyed); Library is oligo-dT primed and directionally  
cloned (EcoRV site is destroyed upon cloning). Average  
insert size 1.42 kb. Library was constructed by  
(Invitrogen). Note: this is a NIH\_MGC Library."  
BASE COUNT 159 a 288 c 247 g 156 t 2 others  
ORIGIN

Alignment Scores:  
Pred. No.: 2.42e-63 Length: 852  
Score: 1165.50 Matches: 237  
Percent Similarity: 90.49% Conservative: 1  
Best Local Similarity: 90.11% Mismatches: 8  
Query Match: 80.38% Indels: 17  
Gaps: 3

SEQ4 (1-266) x CD516283 (1-852)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
|||||  
Db 95 CAGAAAGGAGCTGGGGCCCGCCAGGAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCT 154  
|||||

Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
|||||  
Db 155 GAGGACCATGACATGGGGGACCTGGTGTGCTGAGGAGATCTGGTACCCACGACGCGC 214  
|||||

Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60  
|||||  
Db 215 TTCTCTCTGTGCAGACCTGCCCCGGCGCTGGGGCATGGTCCACACGACGACCGCAGC 274  
|||||

Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
|||||  
Db 275 TCATCTACCAAGAGTGGCGTGGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 334  
|||||

Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
|||||  
Db 335 GCCCCCAGGCTCCACTGGCACCCCTCCGAGGGGCTGGCTCCGATGATTTGATGGTGAC 394  
|||||

Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
|||||  
Db 395 CTGGGAATGGGGGAGCAAGGGGCTCAAAAGCCCTCCACACATGACCCCGCCCTCTA 454  
|||||

Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
|||||  
Db 455 CAGCGGTACAGTGGAGGACCCACAGTACCCCTGCGCCCTGAGACTGATGGCTAGCTTGGC 514  
|||||

Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
|||||

Db 515 CCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACCCAGCAGATGTTGGCCCCCAGCCC 574  
|||||

Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
|||||

Db 575 CCTTGGCCCCGAGAGGGCCCTCTGCCTGCTGCCGAGCTGCTGTGTCACCTCTGGAAGG 634  
|||||

Qy 181 ProlysThrLeuSerPro-GlyLysAsnGlyValValLysAspValPheAlaPheGlyG1 200  
|||||

Db 635 GCCAAGACTCTCTCCCGAGGNAAGATGGGTGCTCAAGACCGCTTTTGCTTGGGG 694  
|||||

Qy 200 yAlaValGluAsnProGluTyrLeuThr-ProGlnGlyGlyAla-AlaProGlnProHis 219  
|||||

Db 695 TGCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGCTTGCCTCAGCCCCAC 754  
|||||

Qy 220 -ProProProAla-PheSerProAla---PheAspAsnLeuTyrTyrTrpAspGlnAsp 238  
|||||

Db 755 CCCTCTCTCTGCCNTTTCAGCCCGAGCCCTTTTCGACAAC-----C 793  
|||||

Qy 238 roPro-----GluArgGlyAlaProProSerThrPheLysGlyThrProThr 253  
|||||

Db 794 CTCCTATTAACTGGGGAACAGGAGCCACCCCAAGAGCGGGGGGCTTCCAACC 850  
|||||

RESULT 14  
CA487981

LOCUS CA487981 932 bp mRNA linear EST 14-NOV-2002  
DEFINITION AGENCOURT\_10808060 MAPcL Homo sapiens cDNA clone IMAGE:6719711 5',  
mRNA sequence.  
ACCESSION CA487981  
VERSION CA487981.1 GI:24948694  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 932)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Kristi A. Eglund, Ira Pastan  
cDNA Library Preparation: Invitrogen Corp  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
plate: LLAM14278 row: a column: 23  
High quality sequence stop: 567.  
Location/Qualifiers  
1. 932  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6719711"  
/cell\_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, HTERT-HME1  
, LNCap"  
/lab\_host="EMDH10B"  
/clone\_lib="MAPcL"  
/note="Vector: pCMV-SPORT6; Site\_1: EcoRV; Site\_2: Not I;  
Subtracted with brain, liver, lung, kidney and muscle.  
Directionally cloned. Priming method: oligo-dr. Average  
insert size: 1800 bp. Library amplification: 26,000 fold.  
Kristi A. Eglund, James J. Vincent, Robert Strausberg,  
Bungkook Lee & Ira Pastan: Discovery of new breast  
cancer genes encoding membrane and secreted proteins.  
Manuscript submitted."

BASE COUNT 173 a 328 c 268 g 162 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1.67e-61 Length: 932

Score:	1136.50	Matches:	229
Percent Similarity:	86.03%	Conservative:	5
Best Local Similarity:	84.19%	Mismatches:	20
Query Match:	78.38%	Indels:	18
DB:	14	Gaps:	5
SEQ4 (1-266) x CA487981 (1-932)			
QY	1	GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu	20
DB	67	CAGAATGAGGACTTGGGCCAGCCAGTCCCTGGACAGCACCTTCTACCGCTCAGTCTG	126
QY	21	GluAspAspMetGlyAspLeuValAlaGluTyrLeuValProGlnGlnGly	40
DB	127	GAGGACGATGACATGGGGACCTGTGTGATCTGTAGGAGTATCTGTATCCCGAGGAGGC	186
QY	41	PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer	60
DB	187	TTCTTCTGTCAGACCTCCCGGGCGCTGGGGCATGTGTCACACAGGACCCGAGC	246
QY	61	SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu	80
DB	247	TCATCTACAGGAGTGGCGTGGGACCTGACACTAGGCTGAGGCTCTGAGGAGGAG	306
QY	81	AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp	100
DB	307	GGCCCCAGGTCTCCACTGGCACCTCCGAAGGGCTGGCTCCGATGATTTGATGTGAC	366
QY	101	LeuGlyMetClyAlaAlaGlyLeuGlnSerLeuProThrHisAspProSerProLeu	120
DB	367	CTGGGAATGGGGACGCCAAGGGCTGCAAGGCTCCACACATGACCCAGCCCTCTA	426
QY	121	GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla	140
DB	427	CACGGTACAGTGAAGACCCACAGTACCCCTGCTGAGACTGATGCTGCTGCTG	486
QY	141	ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro	160
DB	487	CCCTGACCTGCAGCCCCAGCCTGAATATGTAACACAGCAGATGTTCCGCCCGCCG	546
QY	161	ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg	180
DB	547	CTTCGCCCGGAGAGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	606
QY	181	ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPhe	200
DB	607	CCCAACACTCTCCCCAGGAGAGATGGGTGCTCAAGACGTTTTTCCCTTTGGGGG	666
QY	200	YAlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHis	219
DB	667	TGCGGGGAGAACCCCGGACTTGGACCCCGGAGGAGAGCTGGCCCTCAGCCCCAC	726
QY	220	Pro-----ProProAlaPheSerProAlaPheAspAsnLeuTyrTyr	233
DB	727	CTCTCTCTGCGCTTAGGCCCAACCTTCTAACAACCT-----TCTATTTACTGG	777
QY	234	TrpAspGlnAspProGlu---ArgGlyAla-----ProProSerThrPhe-----	248
DB	778	GGACCCGGGAGCCCGCCACAGCGGGGGGGGCTCCACCCCGGAGCCACATTTTATGA	837
QY	249	---LysGlyThrProThrAlaGluAsnPro	257
DB	838	GGGAGCCCTTACCGGCCCGAGATAAACCC	867
RESULT 15			
AUI23871			
LOCUS	AUI23871	685 bp	mRNA
DEFINITION	NT2RM2 Homo sapiens cDNA clone NT2RM2001211 5', mRNA		
ACCESSION	sequence.		
VERSION	AUI23871		
KEYWORDS	AUI23871.1 GI:10948587		
SOURCE	EST.		
	Homo sapiens (human)		

ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 685) Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and Isogai,T.		
TITLE	HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S., Isogai,T.)		
JOURNAL	Unpublished		
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.		
FEATURES	Location/Qualifiers		
source	1..685 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="NT2RM2001211" /cell_type="teratocarcinoma" /cell_line="NT2" /clone_lib="NT2RM2" /note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal precursor cells"		
BASE COUNT	137 a	224 c	198 g 123 t 3 others
ORIGIN			
Alignment Scores:			
Pred. No.:	4.86e-61	Length:	685
Score:	1127.00	Matches:	213
Percent Similarity:	96.83%	Conservative:	1
Best Local Similarity:	96.38%	Mismatches:	6
Query Match:	77.72%	Indels:	1
DB:	9	Gaps:	0
SEQ4 (1-266) x AUI23871 (1-685)			
QY	1	GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu	20
Db	11	AAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTCTG	70
QY	21	GluAspAspMetGlyAspLeuValAlaGluTyrLeuValProGlnGlnGly	40
Db	71	GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTGTACCCAGCAGGC	130
QY	41	PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer	60
Db	131	TTCTTCTGTCAGACCTCCCGGGCGCTGGGGCATGTCCACACAGGACCCGAGC	190
QY	61	SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu	80
Db	191	TCATCTACGAGGAGTGGCGTGGGACCTGACACTAGGCTGAGGCTCTGAAGAGGAG	250
QY	81	AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp	100
Db	251	GGCCCCAGGTCTCCACTGGCACCTCCGGAAGGGCTGGCTCCGATGATTTGATGTGAC	310
QY	101	LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu	120
Db	311	CTGGGAATGGGGACGCCAAGGGCTGCAAGGCTCCCGACACATGACCCCGAGCCCTCTA	370
QY	121	GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla	140
Db	371	CAGCGGTACAGTGAAGACCCACAGTACCCCTGCCCTCTGAGACTGATGGCTGCTG	430

QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
 DB |||||||  
 QY 431 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACCCAGCCAGATGTTGGCCCCCAGCCC 490  
 DB |||||||  
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
 DB |||||||  
 QY 491 CCTTCGCCCGAGAGGGCCCTGCTGCTGCCGACCTGCTGGTGCCACTCTGAAAGG 550  
 DB |||||||  
 QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
 DB |||||||  
 QY 551 GCCAAGACTCTCCCCAGGAGAGAAATGGGGTCTCAAAAGACGTTTTCCTTTGGGGGT 610  
 DB |||||||  
 QY 201 AlaValGluAsnProGluTyr-LeuThrProGlnGlyAlaAlaProGlnProHisPr 220  
 DB |||||||  
 QY 611 GCCGTGGAGAACCCCGAGTACTTTTGACACCCCCANGGAGGAACCTGCCCCCTTAAGCCACC 670  
 DB |||||||  
 QY 220 O 220  
 DB I  
 QY 671 C 671

Search completed: October 16, 2003, 17:04:49  
 Job time : 2095.84 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 16, 2003, 11:08:07 ; Search time 274.965 Seconds  
(without alignments)  
2540.503 Million cell updates/sec

Title: SEQ4

Perfect score: 1450

Sequence: 1 QNEDIGPASPLDSTFYRSL.....TFKGTPTAENPEYLGLDVVP 266

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-DB=PublishedApplications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=BLOSUM62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=HOLLERAN480 -ECGN\_1\_1\_397 @runat.15102003.131915\_20662  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PublishedApplications\_NA.\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1450	100.0	1713	12	US-10-378-393-14
					Sequence 14, Appl

2	1450	100.0	1755	10	US-09-930-125-6
3	1450	100.0	1767	10	US-09-930-125-4
4	1450	100.0	1773	10	US-09-930-125-7
5	1450	100.0	2411	12	US-10-378-393-10
6	1450	100.0	3768	10	US-09-854-356-9
7	1450	100.0	3768	10	US-09-930-125-1
8	1450	100.0	3768	12	US-10-313-644-1
9	1450	100.0	4473	11	US-09-441-411-5
10	1450	100.0	4473	12	US-10-101-510-81
11	1450	100.0	4473	14	US-10-146-473-32
12	1450	100.0	4473	14	US-10-207-655-44
13	1450	100.0	4606	12	US-09-971-392-70
14	1442	99.4	1806	10	US-09-930-125-5
15	1442	99.4	3765	12	US-10-207-498-5
16	1442	99.4	3768	9	US-09-811-123-8
17	1442	99.4	3768	9	US-09-811-115-2
18	1442	99.4	4530	10	US-09-877-172-11
19	1442	99.4	4530	12	US-10-007-926A-119
20	1442	99.4	4530	12	US-10-101-510-124
21	1442	99.4	4530	12	US-10-338-730-1
22	1442	99.4	4530	14	US-10-177-293-125
23	1442	99.4	4543	10	US-09-769-508-1
24	1442	99.4	4642	14	US-10-198-846-10896
25	1442	99.4	9274	9	US-09-811-123-7
26	1442	99.4	9274	9	US-09-811-115-1
27	1208	83.3	3955	10	US-09-870-759-117
28	1208	83.3	3955	10	US-09-854-356-10
29	1208	83.3	3955	12	US-09-751-708A-117
30	1205	83.1	3771	10	US-09-854-356-11
31	1183	81.6	1665	9	US-09-821-883-6
32	1183	81.6	2070	9	US-09-821-883-7
33	1177	81.2	651	9	US-09-821-883-26
34	1177	81.2	1191	9	US-09-821-883-28
35	1177	81.2	1692	9	US-09-821-883-8
36	1177	81.2	2091	9	US-09-821-883-9
37	1177	81.2	3537	9	US-09-821-883-30
38	1041.5	71.8	1115	14	US-10-102-806-165
39	790.5	54.5	507	11	US-09-911-904-116
40	674	46.5	452	11	US-09-918-995-32765
41	647.5	44.7	1046	9	US-09-925-301-390
42	555	38.3	455	11	US-09-918-995-14628
43	475.5	33.1	343	12	US-10-007-926A-118
44	475.5	33.1	343	12	US-10-007-926A-444
45	435	30.0	519	9	US-09-925-301-642

ALIGNMENTS

RESULT 1  
US-10-378-393-14  
; Sequence 14, Application US/10378393  
; Publication No. US20030182668A1  
; GENERAL INFORMATION:  
; APPLICANT: Bol, David K.  
; APPLICANT: Carboni, Joan M.  
; APPLICANT: Rowley, Ronald B.  
; APPLICANT: Wong, Tai W.  
; APPLICANT: Lee, Francis Y.  
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVAT  
; FILE REFERENCE: D0254 NP  
; CURRENT APPLICATION NUMBER: US/10/378,393  
; CURRENT FILING DATE: 2003-03-03  
; PRIOR APPLICATION NUMBER: US 60/360,889  
; PRIOR FILING DATE: 2002-03-01  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 14  
; LENGTH: 1713  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-378-393-14

```

Alignment Scores:
Pred. No.: 5,75e-118 Length: 1713
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ4 (1-266) x US-10-378-393-14 (1-1713)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 913 CAGAATGAGACTGGGCCAGCCAGTCCCTTGGACAGCACCTTACCGCTCACTGCTG 972
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 973 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 1032
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
Db 1033 TTCTTCTGTCAGACCTGCCCGGGCGCTGGGGCATGCTCCACACAGCAGCCGAGC 1092
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 1093 TCATCTACAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 1152
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 1153 GCGCCAGGTCTCCATGGCCAGCCCTCGAAGGGGCTGGCTCCGATGATTTGATGGTGAC 1212
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 1213 CTGGGAATGGGGCAGCAAGGGCTGCAAAAGCTCCCCACACATGACCCCGCCCTCTA 1272
QY 121 GluArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 1273 CAGCGGTACAGTGGAGACCCACAGTACCCCTGCTGAGACTGATGGCTACGTTGCC 1332
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 1333 CCCTTGACCTTGACCCCGCCAGCCCTGATATGTGACCCAGCCAGATGTTGCGCCCGCAGCC 1392
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 1393 CCTTCGCGCCGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGCCTCTGGAAGG 1452
QY 181 ProlysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 1453 CCCAAGACTCTCTCCCGAGGAAGAATGGGGTCTGCAAAAGACGTTTTTGGCTTTGGGGGT 1512
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 1513 GCGGTGGAGAACCCGAGTACTTGCACACCCCGAGGAGACTGCCCTCTAGCCCGCCACCT 1572
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 1573 CCTCCTGCCCTCAGCCCGAGCCTTCGACAACTCTATTACTGGGACAGGACCCAGCAGAG 1632
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 1633 CGGGGGGCTCCACCCAGACCTTTCAAGGGGACACCTACGGCAGAGAACCAGAGTACCTG 1692
QY 261 GlyLeuAspValProVal 266
Db 1693 GGTCTGGACGTGCCAGTG 1710

RESULT 2
US-09-930-125-6
; Sequence 6, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Lodes, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-930-125-6

Alignment Scores:
Pred. No.: 5,9e-118 Length: 1755
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ4 (1-266) x US-09-930-125-6 (1-1755)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 946 CAGAATGAGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTACCGCTCACTGCTG 1005
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 1006 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
Db 1066 TTCTTCTGTCAGACCTGCCCGGGCGCTGGGGCATGCTCCACACAGCAGCCGAGC 1125
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 1126 TCATCTACAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 1185
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 1186 GCGCCCGAGCTCTCCATGGCCAGCCCTCGAAGGGGCTGGCTCCGATGATTTGATGGTGAC 1245
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 1246 CTGGGAATGGGGCAGCAAGGGCTGCAAAAGCTCCCCACACATGACCCCGCCCTCTA 1305
QY 121 GluArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 1306 CAGCGGTACAGTGGAGACCCAGTACCCCTGCCCTCTGAGACTGATGGCTACGTTGCC 1365
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 1366 CCCTGACCTGACGACCCCGCCAGCCTGAATATGTAACAGCAGCATGTTGCGGCCCGCAGCC 1425
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 1426 CCTTCGCGCCGAGAGGGCCCTCTGCTGCTGCCGAGCTGCTGGTGCCTCTGGAAGG 1485
QY 181 ProlysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 1486 CCCAAGACTCTCTCCCGAGGAAGAATGGGGTCTGCAAAAGACGTTTTTGGCTTTGGGGGT 1545
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 1546 GCGGTGGAGAACCCGAGTACTTGCACACCCCGAGGAGACTGCCCTCTAGCCCGCCACCT 1605
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240

```



```

Db      1606 CCTCTGCTTCAGCCAGCCCTTCGACAACTCTATTACTGGAGCAGGACCCACAGAG 1665
QY      241  ArgGlyAlaProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db      1666 CGGGGGCTCCACCCAGCACCTTCAAAGGGACACTACGGCAGAGAACCCAGAGTACCTG 1725
QY      261  GlyLeuAspValProVal 266
Db      1726 GGTCTGGAGCTGCCAGTG 1743

RESULT 3
US-09-930-125-4
; Sequence 4, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedwick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-4

Alignment Scores:
Pred. No.:          5,94e-118      Length:          1767
Score:             1450.00      Matches:          266
Percent Similarity: 100.00%      Conservative:    0
Best Local Similarity: 100.00%      Mismatches:     0
Query Match:       100.00%      Indels:         0
DB:                10          Gaps:         0

SEQ4 (1-266) x US-09-930-125-4 (1-1767)
QY      1  GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db      946 CAGAATGAGGACTTGGGCCCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1005
QY      21  GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db      1006 GAGGACGATGACATGGGGACCTGTGTGATGCTGAGGAGTATCTGTGTACCCAGCAGGCG 1065
QY      41  PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db      1066 TTCTTCTGTCCAGACCTTCCCGGGCGCTGGGGCATGGTCCACACAGCAGCCAGC 1125
QY      61  SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 80
Db      1126 TCATCTACGAGGAGTGGGGGTGGGACCTGACACCTAGGAGCTGGAGCCCTCTCAAGAGGAG 1185
QY      81  AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db      1186 GCCCCAGGCTCTCCACTGGCACCCCTCCGAAGGGGCTGGCTCCGATGTATTGATGGTGAC 1245
QY      101  LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db      1246 CTGGGAATGGGGGACCCAGAGGGGTGCAAAGCCCTCCCCACACATGACCCAGCCCTCTA 1305
QY      121  GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db      1306 CAGCGGTACAGTGGAGGACCCACAGTACCCCTGCCCTCTGAGACTGTAGGCTACGCTTGCC 1365

```

```

QY      141  ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db      1366 CCCCTGACTCGAGCCCCCAGCCTGAATATGTGAACACGACAGATGTTGGGCCCCAGCCC 1425
QY      161  ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db      1426 CCTTCGCCCGAGAGAGGCCCTCTGCTGCTGCCGACCTGCTGGTGCCACTCTGGAAGG 1485
QY      181  ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db      1486 CCCAAGACTCTCTCCCGAGGGAAGATGGGGTCTGCAAGAGAGCTTTTGGCTTTGGGGGT 1545
QY      201  AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db      1546 GCCGTGGAGAACCCCGAGTACTTGACACCCAGGAGGAGCTGCCCTCAGCCCCACCT 1605
QY      221  ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrPheAspGlnAspProGlu 240
Db      1606 CCTCTGCTTTCAGCCAGCCCTTCGACAACTCTATTACTGGAGCAGGACCCACAGAG 1665
QY      241  ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db      1666 CGGGGGGTCTCCACCCAGCACCTTCAAAGGGACACTTACGGCAGAGAACCCAGAGTACCTG 1725
QY      261  GlyLeuAspValProVal 266
Db      1726 GGTCTGGAGCTGCCAGTG 1743

RESULT 4
US-09-930-125-7
; Sequence 7, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedwick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-7

Alignment Scores:
Pred. No.:          5,96e-118      Length:          1773
Score:             1450.00      Matches:          266
Percent Similarity: 100.00%      Conservative:    0
Best Local Similarity: 100.00%      Mismatches:     0
Query Match:       100.00%      Indels:         0
DB:                10          Gaps:         0

SEQ4 (1-266) x US-09-930-125-7 (1-1773)
QY      1  GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db      970 CAGAATGAGGACTTGGGCCCGCCAGCTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1029
QY      21  GluAspAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db      1030 GAGGACGATGACATGGGGACCTGTGTGATGCTGAGGAGTATCTGTGTACCCAGCAGGCG 1089
QY      41  PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60

```

```
Db 1090 TTCCTCTGTCAGACCTGCCCCGGGGCGCTGGGGGCATGGTCCACACAGCAGCCGCGAGC 1149
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 1150 TCATCTACAGAGTGGGGTGGGGACCTGACATAGGGCTGGAGCCCTCTGAAGAGGAG 1209
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 1210 GCCCCAGGTCTCACCTGGCACCTCCGAAGGGCTGGCTCCGATGATTTGATGGTGAC 1269
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 1270 CTGGGAATGGGGCAGCAAGGGCTGCAAAAGCCTCCACACATAGTACCCAGCCCTCTA 1329
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 1330 CAGCGGTACAGTGGAGACCCACAGTACCCCTGCCCCCTGAGACTGATGGCTAGGTTGCC 1389
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 1390 CCCCTGACCTGCAGCCCCCAGCCCTGATATGTGAACCCAGCCAGATGTTGCGCCCCAGCCC 1449
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 1450 CCTTCGCCCCGAGAGGGCCCTCTGCCTGCTGCCGACCTGCTGGTGCCACTCTGGAAGG 1509
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 1510 CCCAAGACTCTCTCCCCAGGGAAGATGGGGTCGTCAAAGACGTTTTTGGCTTTGGGGGT 1569
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 1570 GCCGTGAGAACCCCGAGTACTTGACACCCCGAGGGAGGAGCTGCCCTCAGCCCCACCT 1629
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpaspGlnAspProProGlu 240
Db 1630 CCTCCTGCCCTCAGCCAGCCCTTGACAACCTCTATTACTGGGACCCAGGACCCACCCAGAG 1689
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 1690 CGGGGGGCTCCACCCAGCCTTCAAGGGACACCTACGGCAGAGAACCCAGAGTACCTG 1749
Qy 261 GlyLeuAspValProVal 266
Db 1750 GGTCTGACGTGCCAGTG 1767
```

## RESULT 5

```
US-10-378-393-10
; Sequence 10, Application US/10378393
; Publication No. US2003018268A1
; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVATED
; TITLE OF INVENTION: TYROSINE KINASE RECEPTORS
; FILE REFERENCE: D0254 NP
; CURRENT APPLICATION NUMBER: US/10/378,393
; PRIOR FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,889
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-378-393-10
Alignment Scores:
```

```
Pred. No.: 8.12e-118 Length: 2411
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ4 (1-266) x US-10-378-393-10 (1-2411)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 1611 CAGAATGAGACACTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1670
Qy 21 GlnAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 1671 GAGACCATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCACGAGGC 1730
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
Db 1731 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCATGTGCCACACAGCAGCAGCCGAGC 1790
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 1791 TCATCTACAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 1850
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 1851 GCCCCAGGCTCTCACCTGGCACCTCCGAAGGGCTGGCTCCGATGATTTGATGGTGAC 1910
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 1911 CTGGGAATGGGGCAGCAAGGGCTGCAAAAGCCTCCACACATAGTACCCAGCCCTCTA 1970
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 1971 CAGCGGTACAGTGGAGACCCACAGTACCCCTGCCCTCTGAGACTGATGGCTACGTGGCC 2030
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 2031 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACCCAGCCAGATGTTGCGCCCCAGCCC 2090
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 2091 CCTTCGCCCCGAGAGGGCCCTCTGCCTGCTGCCGACCTGCTGGTGCCACTCTGGAAGG 2150
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 2151 CCCAAGACTCTCTCCCAGGGAAGATGGGGTCTCAAGACGTTTTTGGCTTTGGGGGT 2210
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 2211 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCCCTCAGCCCCACCT 2270
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpaspGlnAspProProGlu 240
Db 2271 CCTCCTGCCCTCAGCCAGCCCTTCGACAACTCTATTACTGGGACCCAGGACCCACAGAG 2330
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 2331 CGGGGGGCTCCACCCAGCACCTTCAAGGGGACACCTACGGCAGAGAACCCAGAGTACCTG 2390
Qy 261 GlyLeuAspValProVal 266
Db 2391 GGTCTGACGTGCCAGTG 2408
```

## RESULT 6

```
US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
```

```
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (1)...(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (1)...(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2026)...(3765)
; OTHER INFORMATION: Intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)...(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)...(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

Alignment Scores:
Pred. No.: 1.28e-117 Length: 3768
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ4 (1-266) x US-09-854-356-9 (1-3768)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db CAGAAATGAGGACTGGGGCCAGCCAGTCCCTTGGACGACACCTCTACCGCTCACTGCTG 3027

Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087

Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60
Db TTCTTCTGTCAGACCTCGCCCGGGCGCTGGGGGCGATGGTCCACACGAGCAGCGCAGC 3147

Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db TCATCTACGAGAGTGGCGGTGGGACCTGACATAGGCGTGGAGCCCTCTGAAGAGAG 3207

Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db GCGCCAGGCTCTCCAGTGGCACCTCCGAAAGGGCTGGCTCGATGATTTGATGGTGAC 3267

Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db CTGGGAATGGGGGAGCCAGAGGGGTGCAAAAGCGCTCCACACATGACCCAGCCCTCTA 3327

Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db CAGCGGTACAGTGAGGACCCACAGTACCCCTGCTGAGCTGATGGCTACGTTGCC 3387

Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db CAGGACGATGATGAGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087

; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (1)...(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (1)...(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2026)...(3765)
; OTHER INFORMATION: Intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)...(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)...(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

Alignment Scores:
Pred. No.: 1.28e-117 Length: 3768
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ4 (1-266) x US-09-854-356-9 (1-3768)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db CAGAAATGAGGACTGGGGCCAGCCAGTCCCTTGGACGACACCTCTACCGCTCACTGCTG 3027

Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087

Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60
Db TTCTTCTGTCAGACCTCGCCCGGGCGCTGGGGGCGATGGTCCACACGAGCAGCGCAGC 3147

Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db TCATCTACGAGAGTGGCGGTGGGACCTGACATAGGCGTGGAGCCCTCTGAAGAGAG 3207

Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db GCGCCAGGCTCTCCAGTGGCACCTCCGAAAGGGCTGGCTCGATGATTTGATGGTGAC 3267

Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db CTGGGAATGGGGGAGCCAGAGGGGTGCAAAAGCGCTCCACACATGACCCAGCCCTCTA 3327

Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db CAGCGGTACAGTGAGGACCCACAGTACCCCTGCTGAGCTGATGGCTACGTTGCC 3387

Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db CAGGACGATGATGAGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087
```



```
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441,411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-5
```

```
Alignment Scores:
Pred. No.: 1 52e-117 Length: 4473
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0
```

SEQ4 (1-266) x US-09-441-411-5 (1-4473)

```
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGATGAGGACTTGGCCAGCCAGTCCCTTGGACAGCACCTTACCGCTCACTGCTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 3202 GAGGACGATGATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
Db 3262 TTCCTCTCCAGACCTTGCCCGGGCGCTGGGGCATGGTCCACACAGCAGCAGCGCAGC 3321
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3322 TCATCTACGAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3381
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3382 GCCCGGAGTCTCCACTGGCACCTTCCGAAAGGGCTGGCTCCGATGATTTGATGGTGAC 3441
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3442 CTGGGAATGGGGGACCCAGAGGGCTGCAAGCCCTCCACACATGACCCCGCCCTCTA 3501
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3502 CAGCGGTACAGTGAAGACCCACAGTACCTCCCTCTGAGACTGATGGCTACGTTGCC 3561
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3562 CCCCTGACTCGAGCCCCAGCCTGAATATGTGAACAGCAGCAGATGTTCCGGCCCCAGGCC 3621
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3622 CCTTCGCCCGAGAGGGCCCTCGCTGCTGCCGACCTGCTGGTGCCACTCTGGAAGG 3681
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3682 CCCAAGACTCTCTCCCGAGGAAGAAATGGGTGCTGCTCAAGACGTTTTTGGCTTTGGGGGT 3741
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 3742 GCCGTGGAGAACCCCGAGTACTTGACACCCAGGAGGAGCTGCCCTTCAGCCCCACCCCT 3801
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 3802 CCTCTGCTTCAGCCCCAGCCTTCGACAACTCTATTACTGGGACGAGCAGCAGCAGAG 3861
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
|||||
```

```
Db 3862 CGGGGGCTCCACCAGCACCTTCAAAGGAGACCTACGGCAGAGAACCCAGAGTACCTG 3921
QY 261 GlyLeuAspValProVal 266
|||||
Db 3922 GGTCTGGACGTGCCAGTG 3939
RESULT 10
US-10-101-510-81
; Sequence 81, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-81
Alignment Scores:
Pred. No.: 1 52e-117 Length: 4473
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
```

SEQ4 (1-266) x US-10-101-510-81 (1-4473)

```
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGATGAGGACTTGGCCAGCCAGTCCCTTGGACAGCACCTTACCGCTCACTGCTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 3202 GAGGACGATGATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60
Db 3262 TTCCTCTCCAGACCTTGGCCAGCCCTGCGGGCATGGTCCACACAGCAGCAGCGCAGC 3321
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3322 TCATCTACGAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3381
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3382 GCCCGGAGTCTCCACTGGCACCTTCCGAAAGGGCTGGCTCCGATGATTTGATGGTGAC 3441
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3442 CTGGGAATGGGGGACCCAGAGGGCTGCAAGCCCTCCACACATGACCCCGCCCTCTA 3501
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3502 CAGCGGTACAGTGAAGACCCACAGTACCTCCCTCTGAGACTGATGGCTACGTTGCC 3561
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3562 CCCCTGACTCGAGCCCCAGCCTGAATATGTGAACAGCAGCAGATGTTCCGGCCCCAGGCC 3621
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3622 CCTTCGCCCGAGAGGGCCCTCGCTGCTGCCGACCTGCTGGTGCCACTCTGGAAGG 3681
```

```
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3682 CCCAAGACTCTCTCCCGAGGAAAGTGGGTGCTCAAGACGCTTTTGGCTTTGGGGT 3741
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3742 GCCGTGGAGAACCCGAGTACTTGGACACCCAGGGAGGAGCTGCCCTCTACGCCCCACCT 3801
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 3802 CCTCCTGCCCTTACGCCAGCCTTCGACAACTCTATTACTGGGACAGGACCCACAGAG 3861
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3862 CGGGGGGCTCCACCCAGCCTTCAAAGGGACACCTACGGCAGAGAACCACAGAGTACCTG 3921
QY 261 GlyLeuAspValProVal 266
Db 3922 GGTCGTGACGTGCCAGTG 3939
RESULT 11
US-10-146-473-32
; Sequence 32, Application US/10146473
; Publication No. US2003010888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-473-32
Alignment Scores:
Pred. No.: 1,52e-117 Length: 4473
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
SEQ4 (1-266) x US-10-146-473-32 (1-4473)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGAATGAGGACTTGGGCCCGACAGCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTGTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 3202 GAGGACGATGACATGGGGGACCTGGTGGATGCTTGAGAGTATCTGTGATCCCGACAGGGC 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60
Db 3262 TTCTTCTGTCACAGCCCTGCCCGGGCGCTGGGGGATGCTCCACAGGACCGACG 3321
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3322 TCATCTACAGGAGTGGCGGTGGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3381
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3381 TTTTCTTCTGTCACAGCCCTGCCCGGGCGCTGGGGGATGCTCCACAGGACCGACG 3441
```

```
Db 3382 GCCCCAGAGTCTCCACTGGCACCCTCCGAAGGGGGTGGCTCCGATGTATTTGATGGTGAC 3441
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3442 CTGGGAATGGGGCAGCCAAAGGGCTGCAAAAGCCTCCACACATGACCCACCCCTCTA 3501
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3502 CAGCGGTACAGTGAGGACCCACAGTACCCCTCTGAGACTGATGGCTACGTTGCC 3561
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3562 CCCTCAGCTGACGCCCGCCAGCCTGAATATGTGAACAGCAGATGTTCGGCCCCAGCCC 3621
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3622 CCTTCGCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGGCTCTGGAAGG 3681
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3682 CCCAAGACTCTCTCCCGAGGAAAGTGGGTGCTCAAGACGCTTTTGGCTTTGGGGT 3741
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 3742 GCCGTGGAGAACCCCGAGTACTTGGACACCCCGAGGAGGAGTCCCTCTCAGCCCCACCT 3801
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 3802 CCTCCTGCCCTTACGCCAGCCTTCGACAACTCTATTACTGGGACAGGACCCACAGAG 3861
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3862 CGGGGGGCTCCACCCAGCAGCCTTCAAAGGGACACCTACGGCAGAGAACCACAGAGTACCTG 3921
QY 261 GlyLeuAspValProVal 266
Db 3922 GGTCGTGACGTGCCAGTG 3939
RESULT 12
US-10-207-655-44
; Sequence 44, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN- IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-207-655-44
Alignment Scores:
Pred. No.: 1,52e-117 Length: 4473
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
SEQ4 (1-266) x US-10-207-655-44 (1-4473)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGAATGAGGACTTGGGCCCGACAGCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTGTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 3202 GAGGACGATGACATGGGGGACCTGGTGGATGCTTGAGAGTATCTGTGATCCCGACAGGGC 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60
Db 3262 TTCTTCTGTCACAGCCCTGCCCGGGCGCTGGGGGATGCTCCACAGGACCGACG 3321
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3322 TCATCTACAGGAGTGGCGGTGGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3381
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3381 TTTTCTTCTGTCACAGCCCTGCCCGGGCGCTGGGGGATGCTCCACAGGACCGACG 3441
```

Db 3202 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3261  
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgSer 60  
 Db 3262 TTCCTCTCTCAGACCCCTGCCCCGGGCGCTGGGGCATGGTCCACACAGCAGCCGCGAGC 3321  
 Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
 Db 3322 TCATCTACAGGAGTGGCGGTGGGACCTGACACTAGGCGTGGAGCCCTCTGAAGAGGAG 3381  
 Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
 Db 3382 GCCCCAGGCTCCACTGGCACCCTCCGAAGGGGCTGGCTCCGATGTATTTGATGGTGAC 3441  
 Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
 Db 3442 CTGGAAATGGGGGAGCCCAAGGGCTGCAAAAGCCTCCCCACATGACACCCAGCCCTCTA 3501  
 Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
 Db 3502 CAGCGGTACAGTGGAGACCCACAGTACCCCTGCCCTCTGAGACTGATGGCTAGCTGCC 3561  
 Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
 Db 3562 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACACAGCCAGATGTTCCGGCCCGC 3621  
 Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
 Db 3622 CCTTGGCCCCAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGCCACCTCTCGAAAGG 3681  
 Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
 Db 3682 CCCAAGACTCTCTCCCGAGGAGAAATGGGGTCTGCAAAAGCGTTTTTGGCGGT 3741  
 Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220  
 Db 3742 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGTGGCCCTTCAGCCCCAC 3801  
 Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTyrAspGlnAspProGlu 240  
 Db 3802 CCTCTGCTTCAGCCCGAGCCTTCGACAACTCTATTACTGGGACGAGCCACAGAG 3861  
 Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
 Db 3862 CGGGGGGCTCCACCAGCACCTTCAAGGGGACCTACGGCAGAGAACCCAGAGTACCTG 3921  
 Qy 261 GlyLeuAspValProVal 266  
 Db 3922 GGTCTGGACGTGCCAGTG 3939

RESULT 13

US-09-971-392-70  
 ; Sequence 70, Application US/09971392  
 ; Publication No. US20030134283A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Peterson, David P.  
 ; APPLICANT: Pearson, Cecelia I.  
 ; APPLICANT: Cocks, Benjamin G.  
 ; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION  
 ; FILE REFERENCE: PA-0029 US  
 ; CURRENT APPLICATION NUMBER: US/09/971.392  
 ; PRIOR FILING DATE: 2001-10-03  
 ; PRIOR APPLICATION NUMBER: 60/237,652  
 ; PRIOR FILING DATE: 2000-10-03  
 ; NUMBER OF SEQ ID NOS: 260  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 70  
 ; LENGTH: 4606  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Template ID: 276948.4

US-09-971-392-70  
 Alignment Scores:  
 Pred. No.: 1.56e-117 Length: 4606  
 Score: 1450.00 Matches: 266  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0  
 SEQ4 (1-266) x US-09-971-392-70 (1-4606)  
 Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 Db 3188 CAGATGAGGACTTGGGCCCGCCAGCTCCCTTGGACAGACACCTTCTACCGCTCACTGCTG 3247  
 Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
 Db 3248 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3307  
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60  
 Db 3308 TTCCTCTCTCAGACCCCTGCCCGGGCGCTGGGGCATGGTCCACACAGGACCCGCGAGC 3367  
 Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
 Db 3368 TCATCTACAGGAGTGGGGTGGGACCTGACACTAGGCGTGGAGCCCTCTGAAGAGGAG 3427  
 Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
 Db 3428 GCCCCAGGCTCCACTGGCACCCTCCGAAGGGGCTGGCTCCGATGTATTTGATGGTGAC 3487  
 Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
 Db 3488 CTGGAAATGGGGGAGCCCAAGGGCTGCAAAAGCCTCCCCACATGACACCCAGCCCTCTA 3547  
 Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
 Db 3548 CAGCGGTACAGTGGAGACCCACAGTACCCCTGCCCTCTGAGACTGATGGTACCTTGC 3607  
 Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
 Db 3608 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACACAGCAGATGTTTCGGCCCCAGCC 3667  
 Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
 Db 3668 CCTTGGCCCCAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGCCACCTCTGGAAGG 3727  
 Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
 Db 3728 CCCAAGACTCTCTCCCGAGGAGAAATGGGGTGGTCAAAAGACGTTTTTTCCTTTTGGGGGT 3787  
 Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220  
 Db 3788 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGTGGCCCTTCAGCCCCACCT 3847  
 Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTyrAspGlnAspProProGlu 240  
 Db 3848 CCTCTGCTTCAGCCCGAGCCTTCGACAACTCTATTACTGGGACGAGACCCACAGAG 3907  
 Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
 Db 3908 CGGGGGGCTCCACCAGCACCTTCAAGGGGACACCTACGGCAGAGAACCCAGAGTACCTG 3967  
 Qy 261 GlyLeuAspValProVal 266  
 Db 3968 GGTCTGGACGTGCCAGTG 3985

RESULT 14

US-09-930-125-5  
 ; Sequence 5, Application US/09930125  
 ; Publication No. US20020193329A1  
 ; GENERAL INFORMATION:

```

; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1806
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-5

Alignment Scores:
Pred. No.: 3,04e-117 Length: 1806
Score: 1442.00 Matches: 265
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.45% Indels: 0
DB: 10 Gaps: 0

SEQ4 (1-266) x US-09-930-125-5 (1-1806)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 1003 CAGAATGAGGACTTGGCCCGCAGCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTGCTG 1062
Qy 21 GluAspAspMetGlyAspLeuValAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 1063 GAGACGATGACATGGGGGACCTGGTGGATGCTGAGAGTATCTGGTACCCAGCAGGCG 1122
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db 1123 TTCTTCTGTCCAGACCTGCCCGGGCGTGGGGCATGTGTCACACAGCAGCCGCGAGC 1182
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 1183 TCATCTACCAAGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 1242
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 1243 GCGCCAGGCTCTCCACTGGGACCCCTCGGAAGGGGCTGGCTCCCATGTATTTGATGGTGAC 1302
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 1303 CTGGGAATGGGGCAGCAGCAAGGGCTGCAAAAGCCTCCCCACACATGACCCCGCCCTCTA 1362
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 1363 CACGGGTACAGTGAAGACCCAGTACCCCTCCCTCTGAGACTGATGGCTACGTTGCC 1422
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 1423 CCCTGACCTGACGCCCGCCAGCCTGAATATGTGAACAGCCAGCATGTTTCGGCCCGACCC 1482
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 1483 CCTTCGCGCCGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGTCGCTGCTGGAAGG 1542
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysaspValPheAlaPheGlyGly 200
Db 1543 GCCAAGACTCTCTCCCGAGGAGAAATGGGGTCTGCTCAAGACAGCTTTTTCGCTTTGGGGT 1602
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 1603 GCGGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCCCTCTACGCCCCACCT 1662

; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1806
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-207-498-5

Alignment Scores:
Pred. No.: 6,38e-117 Length: 3765
Score: 1442.00 Matches: 265
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.45% Indels: 0
DB: 12 Gaps: 0

SEQ4 (1-266) x US-10-207-498-5 (1-3765)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGCCCGCAGCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGAGTATCTGGTACCCAGCAGGCG 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db 3088 TTCTTCTGTCCAGACCTGCCCGGGCGTGGGGCATGTGTCACACAGGACCCGCGAGC 3147
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3148 TCATCTACCAAGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3207
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3208 GCGCCCGAGGCTCTCCACTGGCAGCCCTCCGAAGGGGCTGGCTCCCATGTATTTGATGGTGAC 3267
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3268 CTGGGAATGGGGGAGCAGCAAGGGGCTGCAAAAGCCTCCCGACATGACATGACCCCGCCCTA 3327

; Sequence 5, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; FILE REFERENCE: AND MODULATING INTERACTIONS BETWEEN HEREGULIN AND HER3
; FILE REFERENCE: 30448.103-US-U1
; CURRENT APPLICATION NUMBER: US/10/207,498
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-207-498-5

Alignment Scores:
Pred. No.: 6,38e-117 Length: 3765
Score: 1442.00 Matches: 265
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.45% Indels: 0
DB: 12 Gaps: 0

SEQ4 (1-266) x US-10-207-498-5 (1-3765)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGCCCGCAGCAGTCCCTTGGACAGCACCTTCTACCGCTCAGTGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGAGTATCTGGTACCCAGCAGGCG 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db 3088 TTCTTCTGTCCAGACCTGCCCGGGCGTGGGGCATGTGTCACACAGGACCCGCGAGC 3147
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3148 TCATCTACCAAGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3207
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3208 GCGCCCGAGGCTCTCCACTGGCAGCCCTCCGAAGGGGCTGGCTCCCATGTATTTGATGGTGAC 3267
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3268 CTGGGAATGGGGGAGCAGCAAGGGGCTGCAAAAGCCTCCCGACATGACATGACCCCGCCCTA 3327
```



QY	121	GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla	140
Db	3328	CAGCGGTACAGTGAGACCCACACACTACCCCTGCCCTGTGACATGATGGCTACGTTGCC	3387
QY	141	ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro	160
Db	3388	CCCTTGACCTGACGCCCGCCAGCCTGAATATGTGAACCCAGCCAGATGTTGGGCCCGAGCCC	3447
QY	161	ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrIeuGluArg	180
Db	3448	CCCTCGCCCGGAGAGGGCCCTGTGCTGTGCCCGACCTGTGTGGTCCACTCTGGAAAGG	3507
QY	181	ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly	200
Db	3508	GCCAGACTCTCTCCCGAGGGAAGATGGGGTCTGTCAAAGACGTTTTTTCCTTTGGGGGT	3567
QY	201	AlaValGluAsnProGluTyrIeuThrProGlnGlyGlyAlaAlaProGlnProHisPro	220
Db	3568	GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCCCTCAGCCCCACCCCT	3627
QY	221	ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu	240
Db	3628	CCTCTGCTCTCAGCCCGAGCCTTCGACAACTCTATTACTGGGACCGAGCCACCCAGAG	3687
QY	241	ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu	260
Db	3688	CGGGGGGTCCACCCAGCACCTTCAAGGGGACACCTACGGCAGAGACCCAGAGTACCTG	3747
QY	261	GlyLeuAspValProVal	266
Db	3748	GGTCTGGAGGTGCCAGTG	3765

Search completed: October 17, 2003, 03:54:36  
Job time : 299.965 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 252.648 Seconds  
(without alignments)  
2842.104 Million cell updates/sec

Title: SEQ4  
Perfect score: 1450  
Sequence: 1 QNEDLGASPLDSTFYRSL.....TFKGTPTAENPEYLGLDVPV 266

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/HOLLERAN480/runat\_15102003\_131912\_20526/app\_query.fasta\_1.4685  
-DB=N\_Geneseq\_19Jun03 -QMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 -CGN\_1\_1\_1758 -runat\_15102003\_131912\_20526 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGESQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1450	100.0	1755	24	AAD32746	Human CDNA for the
2	1450	100.0	1767	24	AAD32744	Human CDNA for the
3	1450	100.0	1773	24	AAD32747	Human CDNA for the
4	1450	100.0	3768	17	AA740739	HER-2/neu oncogene
5	1450	100.0	3768	20	AAH01912	Human HER-2/neu on
6	1450	100.0	3768	22	AAH23392	Human HER-2/neu pr
7	1450	100.0	3768	24	AAD32743	Human Her-2/neu pr
8	1450	100.0	3768	24	ABA92250	Human Her-2/neu CD
9	1450	100.0	3768	24	ABK10730	Human Her-2/neu DN
10	1450	100.0	4299	14	AAQ46083	Sequence encoding
11	1450	100.0	4472	21	AAA14812	CDNA encoding the
12	1450	100.0	4473	19	ABQ76220	Human tumour anti
13	1450	100.0	4473	20	AAZ31071	HER-2 nucleic acid
14	1450	100.0	4473	24	ABZ34969	Human gene express
15	1450	100.0	4473	24	ABZ34904	Human Her-2 DNA.
16	1442	99.4	1806	24	AAD32745	Human CDNA for the
17	1442	99.4	3678	24	ABK86207	CDNA encoding huma
18	1442	99.4	3768	21	AAA09455	Human heregulin 2
19	1442	99.4	3768	24	ABZ35744	Human ERBB2 polyu
20	1442	99.4	3768	24	ABX09987	Human ERBB2 DNA fr
21	1442	99.4	3768	24	AAD43935	Human HER-2 CDNA.
22	1442	99.4	3768	24	AAD43986	Human Her2 antigen
23	1442	99.4	3768	24	ABV78168	Human ERBB2 DNA SE
24	1442	99.4	3768	24	ABL91709	Human polynucleoti
25	1442	99.4	3768	24	ABK14058	Human HER2 (Erbb2)
26	1442	99.4	4530	16	AA701585	Her-2/neu (ERBB2/c
27	1442	99.4	4530	18	AA71253	Human HER2 gene.
28	1442	99.4	4530	21	AAZ60815	Nucleotide sequenc
29	1442	99.4	4530	22	AAD19731	Human tyrosine kin
30	1442	99.4	4530	24	ABZ35012	Human gene express
31	1442	99.4	4530	24	ABV94128	Breast carcinoma r
32	1442	99.4	4530	24	ABN85585	Human HER2/neu SEQ
33	1442	99.4	4530	24	ABK83918	Human CDNA differe
34	1442	99.4	4530	25	ACC50139	Breast cancer asso
35	1442	99.4	4530	25	ABQ83856	Human Her2/Neu enc
36	1442	99.4	9274	24	AAD43934	HER-2 transgene pl
37	1442	99.4	9274	24	ABK14057	Human HER2 (Erbb2)
38	1435	99.0	9274	22	AAF24297	HER2 transgene pla
39	1208	83.3	3955	16	AA701590	Rat neu promoter.
40	1208	83.3	3955	21	AAA89753	Rat HER-2/neu prot
41	1205	83.1	2125	19	AAV21727	Humanised vector p
42	1205	83.1	2763	24	ABA92252	Mouse Her-2/neu ex
43	1205	83.1	2781	24	ABA92253	Her-2/neu extracel
44	1205	83.1	3771	21	AAA89737	Mouse Her-2/neu CD
45	1205	83.1	3771	22	AAH42210	Nucleotide sequenc

ALIGNMENTS

RESULT 1

AAD32746

ID AAD32746 standard; CDNA; 1755 BP.

AC AAD32746;

DT 01-JUL-2002 (first entry)

DE Human CDNA for the clone HICD\_native\_coding\_region.

XX Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;  
KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.  
XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1755

```

FT  /**tag= a
FT  /product= "Human protein encoded by cDNA for the clone
FT  HICD_native_coding_region"
FT  /transl_except= (pos:1741..1752, aa:Leu-Glu)
FT  /note= "CDS does not include stop codon"
FT  /partial
XX
XX  WO200214503-A2.
XX
XX  21-FEB-2002.
XX
XX  14-AUG-2001; 2001WO-US41733.
XX
XX  14-AUG-2000; 2000US-225152P.
XX  28-SEP-2000; 2000US-236428P.
XX  21-FEB-2001; 2001US-270520P.
XX
XX  (CORI-) CORIXA CORP.
XX
XX  Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;
XX  Mcneill PD, Vedvick TS;
XX
XX  WPI; 2002-280758/32.
XX  P-PSDB; RAE20483.
XX
XX  Novel isolated Her-2/Neu polypeptide composition useful for therapy,
XX  prevention and diagnosis of cancer, preferably breast cancer -
XX
XX  Example 5; Page 118-119; 129pp; English.
XX
XX  The invention relates to an isolated Her-2/Neu polypeptide composition
XX  effective for eliciting an immune response. The invention is useful for
XX  eliciting an immune response in a patient, where the patient is human
XX  leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
XX  The composition is useful for the therapy and diagnosis of cancer.
XX  preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
XX  and other compositions for the diagnosis, prevention and treatment of
XX  human malignancies, for stimulating and/or expanding T cells specific for
XX  Her-2/Neu polypeptide and for inhibiting the development of cancer in a
XX  patient. The invention is useful for stimulating a T cell response in a
XX  human patient, as probe or primer for nucleic acid hybridisation, to
XX  selectively form duplex molecules with complementary stretches of the
XX  entire Her-2/Neu gene or gene fragments of interest, to isolate a full
XX  length gene from a suitable library, and to direct expression of a
XX  polypeptide in appropriate host cells. The composition is useful in
XX  prophylactic or therapeutic applications and for the treatment of cancer,
XX  preferably for the immunotherapy of breast cancer and other Her-2/Neu-
XX  associated malignancies. The invention is useful in gene therapy. The
XX  present sequence is human cDNA for the clone HICD_native_coding_region.
XX
XX  SQ  Sequence 1755 BP; 376 A; 517 C; 531 G; 331 T; 0 other;

Alignment Scores:
Pred No.: 4.69e-67 Length: 1755
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

SEQ4 (1-266) x AAD32746 (1-1755)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 946 CAGAATGAGGACITGGGCCCCAGCAGTCCCTTGGACAGACCTTCTACCGCTCACTGCTG 1005
QY 21 GlnAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Dy 1006 GAGCAGCATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGC 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyClyMetValHisArgHisArgSer 60
Dy 1066 TTCTTCTGTCCAGACCTGCCCCGGGGCGCTGGGGGGCATGTGTCCACCAGCAGCCGAGC 1125

```

```

QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Dy 1126 TCATCTACCAAGAGTGGCGGTGGGAGCTGACACTAGGCTGGAGCCCTCTGAAGAGGAG 1185
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Dy 1186 GCCCCAGGTTCCACATGGCACCCCTCCGAAAGGGCTGGCTCCGATGATTTGATGGTGAC 1245
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Dy 1246 CTGGGAATGGGGGAGCCCAAGGGGTGCAAAAGCTTCCCCACATGAGCCCGAGCCCTCTA 1305
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Dy 1306 CAGCGGTACAGTGAGGACCCACAGTACCCCTCGCCCTCTGAGACTGATGGCTAGCTGCC 1365
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Dy 1366 CCCCTGACCTGCAGCCGCCAGCCTGAATATGTGAACCCAGCCAGATGTTGGCCGCCAGCC 1425
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Dy 1426 CCTTCGCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGCCACTCTGGAAGG 1485
QY 181 ProlysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Dy 1486 CCCAAGACTCTCTCCCCAGGGAAGAATGGGGTCTCAAAAGACGTTTTTGCCTTTGGGGGT 1545
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Dy 1546 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCGCCCTCAGCCGCCACCT 1605
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
Dy 1606 CCTCCTCGCTTCAGCCCGACCTTCGACAACCTCTATTACTGGGACCCAGAGCCACCCAGAG 1665
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Dy 1666 CGGGGGGCTCCACCCAGCACCCTTCAAAAGGAGACCTTACGGCAGAGAACCCAGAGTACCTG 1725
QY 261 GlyLeuAspValProVal 266
Dy 1726 GGTCTGGACGTGCCAGTG 1743
XX
XX  RESULT 2
XX  AAD32744
XX  ID AAD32744 standard; cDNA; 1767 BP.
XX
XX  AAD32744;
XX
XX  01-JUL-2002 (first entry)
XX
XX  Human cDNA for the clone HICD_CT_His_coding_region.
XX
XX  Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
XX  human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.
XX
XX  Homo sapiens.
XX
XX  Key Location/Qualifiers
XX  CDS 1..1764
XX  /tag= a
XX  /product= "Human protein encoded by cDNA for the clone
XX  HICD_CT_His_coding_region"
XX
XX  WO200214503-A2.
XX
XX  21-FEB-2002.
XX
XX  14-AUG-2001; 2001WO-US41733.
XX
XX  14-AUG-2000; 2000US-225152P.
XX
XX  PR

```

PR 28-SEP-2000; 2000US-236428P.  
PR 21-FEB-2001; 2001US-270520P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;  
PI McNeill PD, Vedvick TS;  
XX  
XX WPI: 2002-280758/32.  
DR P-PSDB; AAE20481.  
XX  
PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,  
PT prevention and diagnosis of cancer, preferably breast cancer -  
XX  
XX Example 5; Page 117-118; 129pp; English.  
XX  
XX The invention relates to an isolated Her-2/Neu polypeptide composition  
CC effective for eliciting an immune response. The invention is useful for  
CC eliciting an immune response in a patient, where the patient is human  
CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.  
CC The composition is useful for the therapy and diagnosis of cancer,  
CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine  
CC and other compositions for the diagnosis, prevention and treatment of  
CC human malignancies, for stimulating and/or expanding T cells specific for  
CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a  
CC patient. The invention is useful for stimulating a T cell response in a  
CC human patient, as probe or primer for nucleic acid hybridisation, to  
CC selectively form duplex molecules with complementary stretches of the  
CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full  
CC length gene from a suitable library, and to direct expression of a  
CC polypeptide in appropriate host cells. The composition is useful in  
CC prophylactic or therapeutic applications and for the treatment of cancer,  
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-  
CC associated malignancies. The invention is useful in gene therapy. The  
CC present sequence is human cDNA for the clone HICD-CT.His\_coding\_region.  
XX  
SQ Sequence 1767 BP; 381 A; 521 C; 529 G; 336 T; 0 other;

Alignment Scores:  
Pred. No.: 4, 72e-67 Length: 1767  
Score: 1450.00 Matches: 266  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 24

SEQ4 (1-266) x AAD32744 (1-1767)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
DB 946 CAGAAATGAGGACTGGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1005  
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
DB 1006 GAGGACGATGACATGGGGGACCTGGTGATCTGAGGAGTATCTGGTACCCAGCAGGGC 1065  
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60  
DB 1066 TTCCTTCCTCAGACCTGGCCCGGGGCGCTGGGGGCGCTGGCCACAGCAGCCGCGAGC 1125  
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
DB 1126 TCATCTACGAGGAGTGGGGTGGGACCTGACATAGGGCTGGAGCCCTCTGAGAGGAG 1185  
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
DB 1186 GCCCCAGGCTCCACTGGCACCCTCCGAGGGGCTGGCTCCGATGATTTGATGGTGAC 1245  
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
DB 1246 CTGGGAATGGGGGAGCCCAAGGGCTGCAAAAGCCTCCCCACACATGACCCAGCCCTCTA 1305  
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140

DB 1306 CAGCGGTACAGTGAAGACCCACAGTACCCCTCTGAGACTGATGGCTACGTTGCC 1365  
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
DB 1366 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACACAGCCAGATGTTTCGGCCCCAGGCC 1425  
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaValThrLeuGluArg 180  
DB 1426 CCTTCGCCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGGCCACACTCTGGAAGG 1485  
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
DB 1486 CCCAAGACTCTCTCCCCAGGCAAGATGGGCTCGTCAAAGAGCTTTTTCCTTTGGGGGT 1545  
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220  
DB 1546 GCCGTGGAGAACCCCGAGTACTTGACACCCAGGAGGAGGTGCCCTTCAGCCCAACCT 1605  
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240  
DB 1606 CCTCTGCTTTCAGCCAGCCTTCGACAACTCTATTACTGGGACGAGGCCACACAGAG 1665  
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
DB 1666 CGGGGGGCTCCACCCAGCACCTTCAAAGGGACACCTACGGCAGAGACCCAGACTACCTG 1725  
QY 261 GlyLeuAspValProVal 266  
DB 1726 GGTCTGGACGTGCCAGTG 1743  
RESULT 3  
AAD32747  
ID AAD32747 standard; cDNA; 1773 BP.  
XX  
XX AAD32747;  
DT 01-JUL-2002 (first entry)  
DE Human cDNA for the clone HICD\_in\_pPDM\_coding\_sequence.  
KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;  
KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 1..1770  
FT /tag= a  
FT /product= "Human protein encoded by cDNA for the clone  
FT HICD\_in\_pPDM\_coding\_sequence"  
XX  
PN WO200214503-A2.  
PD 21-FEB-2002.  
XX  
XX 14-AUG-2001; 2001WO-US41733.  
XX  
XX 14-AUG-2000; 2000US-225152P.  
PR 28-SEP-2000; 2000US-236428P.  
PR 21-FEB-2001; 2001US-270520P.  
XX  
XX (CORI-) CORIXA CORP.  
XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;  
PI McNeill PD, Vedvick TS;  
XX  
XX WPI: 2002-280758/32.  
DR P-PSDB; AAE20481.  
XX  
XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,  
XX prevention and diagnosis of cancer, preferably breast cancer -  
XX

Example 5; Page 119; 129pp; English.

PS The invention relates to an isolated Her-2/Neu polypeptide composition  
XX effective for eliciting an immune response. The invention is useful for  
CC eliciting an immune response in a patient, where the patient is human  
CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.  
CC The composition is useful for the therapy and diagnosis of cancer.  
CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine  
CC and other compositions for the diagnosis, prevention and treatment of  
CC human malignancies, for stimulating and/or expanding T cells specific for  
CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a  
CC patient. The invention is useful for stimulating a T cell response in a  
CC human patient, as probe or primer for nucleic acid hybridisation, to  
CC selectively form duplex molecules with complementary stretches of the  
CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full  
CC length gene from a suitable library, and to direct expression of a  
CC polypeptide in appropriate host cells. The composition is useful in  
CC prophylactic or therapeutic applications and for the treatment of cancer,  
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-  
CC associated malignancies. The invention is useful in gene therapy. The  
CC present sequence is human cDNA for the clone HICD\_in\_ppdm\_coding\_  
XX sequence.

SQ Sequence 1773 BP; 383 A; 528 C; 530 G; 332 T; 0 other;

Alignment Scores:  
Pred. No.: 4.73e-67 Length: 1773  
Score: 1450.00 Matches: 266  
Percent Similarity: 100.00% Conservativity: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0

SEQ4 (1-266) x AAD32747 (1-1773)

QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
DB 970 CAGAATGAGGACTGGGCCAGCAGCTCCCTTGACAGCAGCCTTCTACCGCTCAGTGCTG 1029  
QY 21 GluAspAspMetGlyAspLeuValAlaGluGlyTyrLeuValProGlnGlnGly 40  
DB 1030 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGAGTATCTGGTACCCCGCAGGCG 1089  
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60  
DB 1090 TTCTTCTGTCAGACCTGCGCGGCGCTGGGGCATGTGTCCACACAGCAGCCGACG 1149  
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
DB 1150 TCATCTACACGAGTGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 1209  
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
DB 1210 GCCCCAGGCTCTCAGCTGGACCCCTCGAAGAGGGCTGGCTCCGATGATTTGATGGTGAC 1269  
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
DB 1270 CTGGGAATGGGCGAGCCAGAGGGCTGCAAGAGCCTCCCCACACATGACCCCGCCCTCTA 1329  
QY 121 GluArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
DB 1330 CAGCGGTACAGTGGAGCCAGTACCCCTGCTGAGACTGATGGCTGATGGTGGCC 1389  
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
DB 1390 CCCTGACCTGAGCCCGCCAGCTGAATGTAACAGCAGGATGTGGCCCGCCAGCC 1449  
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
DB 1450 CCTTGGCCCGAGAGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1509  
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200

Db 1510 CCCAAGACTCTCTCCCGAGGAGAAATGGGGTCTGTCAAAGACGCTTTTGGCTTTGGGGGT 1569  
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220  
Db 1570 GCGGTGGAGAACCCCGAGTACTTGGACACCCCGAGGAGGAGCTGCCCTCAGCCGCCCT 1629  
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240  
Db 1630 CCTCTGCTCTCAGCCCGAGCTTCGACAACTCTATTACTGGGACGAGGACCCAGAG 1689  
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
Db 1690 CGGGGGGCTCCACCCAGACCTTCAAGAGGACACCTACGCGAGAGAACCCAGAGTACCTG 1749  
QY 261 GlyLeuAspValProVal 266  
Db 1750 GGTCTGGACGTGCCAGTG 1767  
RESULT 4  
AAT40739  
ID AAT40739 standard; cDNA; 3768 BP.  
XX  
AC AAT40739;  
XX  
DT 01-JAN-1997 (first entry)  
XX  
DE HER-2/neu oncogene.  
XX  
KW HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;  
KW breast cancer; ovary cancer; colon cancer; lung cancer;  
KW prostate cancer; genetic immunisation; tumour; vaccine; vector;  
KW ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..3765  
FT FT /\*tag= b  
FT FT /note= "nucleotides 2026-3765 (claim 1) code for  
FT FT HER-2/neu intracellular domain"  
XX  
PN W09630514-A1.  
XX  
XX 03-OCT-1996.  
XX  
XX 28-MAR-1996; 96WO-US01689.  
XX  
XX 31-MAR-1995; 95US-0414417.  
XX  
XX (UNIW ) UNIV WASHINGTON.  
XX  
XX Cheever MA, Disis ML;  
XX  
XX WPI; 1996-455361/45.  
XX  
XX P-PSDB; AAW01111.  
XX  
XX DNA encoding HER-2-neu polypeptide(s) - used for prevention or  
XX treatment of malignancies with which the HER-2/neu oncogene is  
XX associated  
XX  
XX Claim 1; Page 49-56; 71pp; English.  
XX  
XX Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or  
XX c-erbB2) protein (AAW01111). The oncogene is overexpressed in various  
XX cancers, including breast, ovarian, colon, lung and prostate, and  
XX appears to induce malignancies through quantitative mechanisms that  
XX result from increased or deregulated expression of an essentially  
XX normal gene product. Nucleotides 2026-3765 of the cDNA sequence  
XX code for the intracellular domain (lys676-Val1255) of the HER-2/neu  
XX protein, which is useful for immunisation against malignancy.  
XX Nucleic acids can be used to direct expression of the intracellular  
XX domain in transformed host cells, or are used, alone or in a viral  
XX vector, for genetic immunisation of an animal.

```
XX SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
Alignment Scores:
Pred. No.: 9,23e-67 Length: 3768
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

SQ04 (1-266) x AAX01912 (1-3768)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
QY 21 GluAspAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGTACCCAGCAGGGC 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60
Db 3088 TTCCTTCTGCACAGCCCTGCCCCGGCGCTGGGGCATGGTCCACACAGGACCCGACGC 3147
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3148 TCATCTACCAGGAGTGGCGTGGGACCTGACACTAGGCTGGAGCCCTCTGAAGAGGAG 3207
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3208 GCGCCCGAGGCTCCACTGGCACCCCTCCGAAGGGCTGCGATGTATTGATGGTGAC 3267
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3268 CTGGGAATGGGGGACCCAGGGGCTGCARAGCCTCCACACATGACCCAGCCCTCTA 3327
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3328 CAGCGGTACAGTAGGAGACCCACACAGTACCCCTGCTGAGCTGATGGCTACGTTGCC 3387
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3388 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACACAGCCAGATGTTTCGGCCCCAGCCC 3447
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3448 CCTTCGCCCCGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGCCACTCTGGAAGG 3507
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3508 CCCAAGACTCTCTCCCGAGGGAAGATGGGGTGGTCAAGAGCTTTTTCCTTTGGGGGT 3567
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 3568 GCGGTGGAGAACCCCGACTACTTGACCCCGAGGAGGAGTGCCTCCCTCAGCCCCACCCCT 3627
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTyrAspGlnAspProProGlu 240
Db 3628 CCTCTGCTCTCAGCCCGACCTTCGACAACTCTATTACTGGACACGAGGCCACACAGAG 3687
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3688 CGGGGGGCTCCACCCAGCACCTTCAAGGGACACCTACGGCAGAGAACCCAGAGTACCTG 3747
QY 261 GlyLeuAspValProVal 266
Db 3748 GGCTCTGGACGTGCCAGTG 3765

RESULT 5
AAX01912
ID AAX01912 standard; DNA; 3768 BP.
XX
```





QY 181 ProLysThrLeuSerProGlyLysAsnGlyValLysAspValPheAlaPheGlyGly 200  
 Db CCCAAGACTCTCTCCCAAGGAAGAAATGGGTCGCTCAAGAGCGTTTTTGGGGGT 3567  
 QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220  
 Db GCGGGGAGAACCCCGAGTACTTGACACCCAGGAGGAGCTGCCCTTCAGCCCAACCT 3627  
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240  
 Db CCTCTGCTTCAGCCCAACCTTCGACAACTCTATTACTGGGACGAGCCACACAGAG 3687  
 QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
 Db CGGGGGGCTCCACCCAGCACCTTCAAGGGGACACCTAGCGCAGAGAACCCACAGAGTACCTG 3747  
 QY 261 GlyLeuaspValProVal 266  
 Db GGTCTGGACGTGCCAGTG 3765  
 RESULT 7  
 AAD32743  
 ID AAD32743 standard; DNA; 3768 BP.  
 AC AAD32743;  
 XX  
 XX 01-JUL-2002 (first entry)  
 XX Human Her-2/neu protein DNA.  
 XX  
 XX Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;  
 KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 1..3768  
 FT /\*tag= a  
 FT /product= "Human Her-2/neu protein"  
 FT 2026..3765  
 FT /\*tag= b  
 FT /note= "Intracellular domain"  
 XX  
 XX WO200214503-A2.  
 XX  
 XX 21-FEB-2002.  
 XX  
 XX 14-AUG-2001; 2001WO-US41733.  
 XX  
 XX 14-AUG-2000; 2000US-225152P.  
 PR 28-SEP-2000; 2000US-236428P.  
 PR 21-FEB-2001; 2001US-270520P.  
 XX  
 XX (CORI-) CORIXA CORP.  
 XX  
 XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;  
 PI McNeill PD, Vedvick TS;  
 XX  
 XX WPI; 2002-280758/32.  
 DR P-PSDB: AAE20479.  
 XX  
 XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,  
 PT prevention and diagnosis of cancer, preferably breast cancer -  
 XX  
 XX Claim 9; Page 109-114; 129pp; English.  
 XX  
 XX The invention relates to an isolated Her-2/Neu polypeptide composition  
 CC effective for eliciting an immune response. The invention is useful for  
 CC eliciting an immune response in a patient, where the patient is human  
 CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.  
 CC The composition is useful for the therapy and diagnosis of cancer,  
 CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine  
 CC and other compositions for the diagnosis, prevention and treatment of

human malignancies, for stimulating and/or expanding T cells specific for  
 Her-2/Neu polypeptide and for inhibiting the development of cancer in a  
 patient. The invention is useful for stimulating a T cell response in a  
 human patient, as probe or primer for nucleic acid hybridisation, to  
 selectively form duplex molecules with complementary stretches of the  
 entire Her-2/Neu gene or gene fragments of interest, to isolate a full  
 length gene from a suitable library, and to direct expression of a  
 polypeptide in appropriate host cells. The composition is useful in  
 prophylactic or therapeutic applications and for the treatment of cancer,  
 preferably for the immunotherapy of breast cancer and other Her-2/Neu-  
 associated malignancies. The invention is useful in gene therapy. The  
 present sequence is human Her-2/neu protein DNA.  
 XX  
 SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 9,23e-67 Length: 3768  
 Score: 1450.00 Matches: 266  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0  
 SEQ4 (1-266) x AAD32743 (1-3768)  
 QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 Db CAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTCTACCGTCTACTGCTG 3027  
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40  
 Db GAGCAGATGACATGGGGACCTGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087  
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 60  
 Db TTCTTCTGTCAGACCTTCCCGGGCGCTGGGGCATGGTCCACACAGCAGCCGCGACG 3147  
 QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
 Db TCATCTACCAGAGTGGGGTGGGACCTGACATGAGGCTGGAGCCCTCTGAAGAGGAG 3207  
 QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
 Db GCGCCAGGTCTCCACTGGCACCTCCGAAAGGGGCTGGCTCCGATGTATTTGATGTGAC 3267  
 QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
 Db CTGGGAATGGGGGAGCCCAAGGGGCTGCAAGAGCTCCCCACACATGACCCCGCCCTCTA 3327  
 QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
 Db CAGCGGTACAGTGAGGACCCACAGTACCCCTCGCCCTCTGAGACTGATGGTACCTGCC 3387  
 QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
 Db CCCCTGACCTGCAGCCCGCCAGCCCTGAATATCTGAACACGAGCAGATGTTGCGCCCGCC 3447  
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
 Db CCTTCCGCCCGGAGAGGGCCCTCTGCTCTCCCGGACCTGTGCTGGTGGCCACTCTGGAAGG 3507  
 QY 181 ProLysThrLeuSerProGlyLysAsnGlyValLysAspValPheAlaPheGlyGly 200  
 Db CCCAAGACTCTCTCCCGGAGGAAGATGGGTCGCTCAAGAGCGTTTTTGGGGGT 3567  
 QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220  
 Db GCGGTGAGAACCCCGAGTACTTGACACCCAGGAGGAGCTGCCCTTCAGCCCAACCT 3627  
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240  
 Db CCTCTGCTTCAGCCCAACCTTCAAGGGGACACCTAGCGCAGAGAACCCACAGAGTACCTG 3687

QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
 Db 3688 CGGGGGCTCCACCCAGCACCTTCAAAGGGACACCTACGGCAGAACCCAGAGTACCTG 3747

QY 261 GlyLeuAspValProVal 266  
 Db 3748 GGTCTGGACGTGCCAGTG 3765

## RESULT 8

ID ABA92250 standard; cDNA; 3768 BP.

XX ABA92250;

DT 17-JUN-2002 (first entry)

DE Human Her-2/neu cDNA.

XX Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;

KW receptor; human; gene therapy; gene; ss.

XX Homo sapiens.

FH Key Location/Qualifiers

FT CDS 1..3768

FT /\*tag= a

FT /product= "Her-2/neu"

XX WO200212341-A2.

XX 14-FEB-2002.

XX 03-AUG-2001; 2001WO-US24283.

XX 03-AUG-2000; 2000US-0632507.

XX (CORI-) CORIXA CORP.

XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX Cheever MA, Gheysen D;

XX WPI; 2002-241743/29.

XX P-PSDB; AAMS1143.

XX Her-2/neu fusion protein for treating or preventing cancer by eliciting  
 PT or enhancing an immune response to the protein, has Her-2/neu  
 PT extracellular domain fused to Her-2/neu intracellular or  
 PT phosphorylation domain

XX Disclosure; Fig 15; 141pp; English.

XX The present sequence is that of human Her-2/neu oncogene cDNA.  
 CC The cDNA encodes Her-2/neu (p185), an oncogenic self protein and  
 CC target for anti-cancer vaccines. The Her-2/neu gene is amplified  
 CC and p185 is overexpressed in a variety of cancers, including breast,  
 CC ovarian, colon, lung and prostate cancer. Her-2/neu is a member  
 CC of the tyrosine kinase family of receptor-like glycoproteins.  
 CC Its overexpression correlates with a poor prognosis in breast and  
 CC ovarian cancers. The invention provides Her-2/neu fusion  
 CC proteins, nucleic acids encoding them, viral vectors, and vaccines  
 CC comprising the fusion proteins or nucleic acid molecules. In  
 CC preferred fusion proteins, the extracellular domain of a Her-2/neu  
 CC protein is fused to a Her-2/neu intracellular domain or  
 CC phosphorylation domain (or its DeltaPD fragment). An immune  
 CC response to Her-2/neu protein is elicited or enhanced by  
 CC administering the fusion protein in the form of a vaccine, or by  
 CC transfecting cells of an animal ex vivo with a nucleic acid  
 CC encoding the fusion protein, and delivering the transfected cells  
 CC to the animal. The fusion proteins, nucleic acids, and isolated  
 CC specific T-cells are useful for inhibiting the development of a  
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer  
 CC in a patient. T cells that specifically react with a Her-2/neu  
 CC fusion protein can be used to remove tumour cells from a sample in

CC order to inhibit the development of cancer in a patient.  
 XX  
 SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

## Alignment Scores:

Pred. No.: 9,23e-67 Length: 3768  
 Score: 1450.00 Matches: 266  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

SEQ4 (1-266) x ABA92250 (1-3768)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 Db 2968 CAGAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTCTACCGCTCAGCTGCTG 3027  
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
 Db 3028 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 3087  
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60  
 Db 3088 TTCTTCTGCCAGACCTGCCCGGGCGCTGGGGGCATGCTCCACAGAGCAGCCGACG 3147  
 QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
 Db 3148 TCATCTACCAGGAGTGGGGACCTGACACTAGGGCTGGAGGCCCTCTGAAGAGGAG 3207  
 QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
 Db 3208 GCCCCAGCTCTCCACTGGCACCTCCGAAGGGGCTGGCTCCGATGTATTGATGGTGAC 3267  
 QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
 Db 3268 CTGGGAATGGGGCAGCAAGGGCTGCAAAAGCTCCCCACACATGACCCAGCCCTCTA 3327  
 QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
 Db 3328 CAGCGGTACGTGAGGACCCACAGTACCCCTCCCTCTGAGACTGATGGCTACGTTGCC 3387  
 QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
 Db 3388 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACCCAGCCAGATGTCGGCCCCAGCCC 3447  
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
 Db 3448 CCTTCGCCCCGAGAGGGGCCCTCTGCTGCTGCCGACCTGCTGGTGCCTCTGGAAGG 3507  
 QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
 Db 3508 CCCAAGACTCTCTCCCCAGGAAGATGGGTGCTCAAGACGCTTTTGGCTTTGGGGGT 3567  
 QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220  
 Db 3568 GCCGTGGACACCCCGAGTACTGTGACACCCAGGAGGAGCTGCCCTCAGCCCCCACCCT 3627  
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240  
 Db 3628 CCTCCTGCCTTCAGCCAGCCCTTCGACAACCTCTATTACTGGGACCCAGGACCCACAGAG 3687  
 QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
 Db 3688 CGGGGGGCTCCACCCAGCACCTTCAAAAGGGACACCTACGGCAGAGAACCAGAGTACCTG 3747  
 QY 261 GlyLeuAspValProVal 266  
 Db 3748 GGTCTGGACGTGCCAGTG 3765

## RESULT 9

ABK10730

ID ABK10730 standard; DNA; 3768 BP.

XX ABK10730;  
 AC 05-JUN-2002 (first entry)  
 DT Human Her-2/neu DNA.  
 DE Human; Her-2/neu; gene; ds; cytostatic; haematological malignancy; CML;  
 XX acute myelogenous leukaemia; AML; chronic myelogenous leukaemia; CLL;  
 KW chronic lymphocytic leukaemia; myeloma; non-Hodgkin's lymphoma; MDS;  
 KW Hodgkin's lymphoma; T cell therapy.  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FT 1..3768  
 FT CDS /\*tag= a  
 FT /product= "Human Her-2/neu"  
 XX  
 XX WO200213847-A2.  
 XX 21-FEB-2002.  
 XX 13-AUG-2001; 2001WO-US25408.  
 XX 14-AUG-2000; 2000US-0638280.  
 PR 28-SEP-2000; 2000US-0675904.  
 XX (CORI-) CORIXA CORP.  
 XX  
 XX Gaiger A, Cheever MA, Hand-zimmermann S;  
 PI  
 XX WPI; 2002-280741/32.  
 DR P-PSDB; AAU77114.  
 XX  
 XX Inhibiting haematological malignancy development by administering  
 PT polypeptide comprising immunogenic portion of Her-2/neu, polynucleotide  
 PT encoding the polypeptide, or antigen presenting cells expressing the  
 PT polypeptide -  
 XX  
 XX Disclosure; Page 66-71; 74pp; English.  
 XX  
 XX The invention relates to a method for inhibiting development of  
 CC haematological malignancy in a patient by administering a polypeptide  
 CC comprising an immunogenic portion of Her-2/neu or a polynucleotide  
 CC encoding the polypeptide. Antigen presenting cells that express the  
 CC protein can also be administered. The sequences are used for inhibiting  
 CC development of haematological malignancy such as acute myelogenous  
 CC leukaemia (AML), chronic myelogenous leukaemia (CML), chronic lymphocytic  
 CC leukaemia (CLL), MDS, myelomas, Hodgkin's lymphoma and non-Hodgkin's  
 CC lymphoma. This sequence represents DNA encoding human Her-2/neu  
 CC polypeptide.  
 XX  
 XX SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;  
 XX  
 XX Alignment Scores:  
 Pred. No.: 9,23e-67 Length: 3768  
 Score: 1450.00 Matches: 266  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0  
 SEQ4 (1-266) x ABK10730 (1-3768)  
 QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 DB 2968 CAGAAATGAGGACTTGGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027  
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnCly 40  
 DB 3028 GAGGACGATGACATGGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 3087

QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60  
 DB 3088 TTCTTCTCTCCAGACCTTCCCGGGCGCTGGGGCGATGGTCCACACAGCAGCCGACG 3147  
 QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGlyProSerGluGlu 80  
 DB 3148 TCATCTACAGAGTGGGGTGGGACCTGACACCTAGGGCTGGAGCCCTCTCAAGAGGAG 3207  
 QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
 DB 3208 GCCCCAGGTCTCCACTGGCACCCTCCGAAGGGGCTGGCTCCGATGATTTGATGGTGAC 3267  
 QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
 DB 3268 GTGGGAATGGGGCAGCCAAAGGGCTGCAAGCCCTCCACACATGACCCAGCCCTCTA 3327  
 QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
 DB 3328 CAGCGGTACAGTGAAGACCCACAGTACCCCTCTGACACTGTATGGCTTACGTGGCTGCC 3387  
 QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
 DB 3388 CCCCTGACCTGCAGCCCCCAGCCCTGAATATGTGAACACAGCAGATGTTGGGCCCCAGCC 3447  
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
 DB 3448 CCTTCGCCCGAGAGGGGCCCTCTGCTGCTGCCGACCTGTGTGGTGCACCTCTGGAAGG 3507  
 QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
 DB 3508 CCCAAGACTCTCTCCCCAGGGAAGTGGGGTCTCAAGAGCTTTTGGCTTTGGGGGT 3567  
 QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220  
 DB 3568 GCCGTGGAGAACCCCGAGTACTTGACACCCAGGAGGAGCTGCCCTCAGCCCCCACCCT 3627  
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240  
 DB 3628 CCTCTGCTTTCAGCCAGCCTTCGACAACCTCTATTACTGGGACCAGGACCACAGAG 3687  
 QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
 DB 3688 CGGGGGGTCTCCACCCAGCACCCTTCAAAGGGACACCTACGGCAGAGAACCCAGAGTACCTG 3747  
 QY 261 GlyLeuAspValProVal 266  
 DB 3748 GGTCTGGACGTGCCAGTG 3765  
 RESULT 10  
 AAQ46083  
 ID AAQ46083 standard; cDNA; 4299 BP.  
 XX  
 AC AAQ46083;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 07-FEB-1994 (first entry)  
 XX  
 DE Sequence encoding a c-erbB-2 tumour antigen.  
 XX  
 KW Tumour antigen; c-erbB-2; glycoprotein; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..4299  
 FT /\*tag= a  
 XX  
 XX WO9316185-A2.  
 XX 19-AUG-1993.  
 XX 05-FEB-1993; 93WO-US01055.  
 PF  
 XX

PR 06-FEB-1992; 92US-0831957.  
 XX (CETU ) CETUS ONCOLOGY CORP.  
 PA (CREA-) CREATIVE BIOMOLECULES INC.  
 XX  
 PI Houston LL, Huston JS, Oppermann H, Ring DB;  
 XX WPI; 1993-272889/34.  
 DR P-PSDB; AAR39568.  
 XX  
 XX New single chain Fv polypeptide binding to C-erbB-2 tumour  
 PT antigen - for imaging or treating breast or ovarian cancer etc.  
 XX  
 XX Disclosure; pages 48-54; 87pp; English.  
 XX  
 CC c-erbB-2 refers to a protein antigen expressed on the surface of  
 CC tumour cells, such as breast and ovarian tumour cells, which is an  
 CC approx. 200,000 mol. wt. acidic glycoprotein having an isoelectric  
 CC pt. of about 5.3 (see AAQ46083, AAR39568).  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 4299 BP; 882 A; 1307 C; 1266 G; 844 T; 0 other;  
 SQ

Alignment Scores:  
 Pred. No.: 1.04e-66 Length: 4299  
 Score: 1450.00 Matches: 266  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 14 Gaps: 0

SEQ4 (1-266) x AAQ46083 (1-4299)

QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 DB CAGAATGAGGACTTGGCCGAGCCAGTCCCTTGGACAGACCTTCTACCGCTCAGTCTG 3027  
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40  
 DB GAGGACGATGACATGGGGACCTGGTGATGCTGAGAGTATCTGGTACCCAGCAGGC 3087  
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 60  
 DB TTCTTCTGTGTCAGACCTGCGCGGGGCTGGGGGATGCTCCACACAGCAGCCGCGAGC 3147  
 QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
 DB TCATCTACAGGAGTGGCGTGGGACCTGACACATAGGGCTGGAGCCCTCTGAAGAGGAG 3207  
 QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
 DB GCCCCAGGTCTCCACTGGCACCTCCGAAGGGGCTGGCTCCGATGTATTTGATGGTGAC 3267  
 QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
 DB CTGGGAATGGGGGAGCCCAAGGGGCTGCAAAAGCTCCCCACATGACCCAGCCCTCTA 3327  
 QY 121 GluArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
 DB CAGCGGTACAGTGGAGACCCACAGTACCTCCCTCTGAGACTGATGGCTACGTGGCC 3387  
 QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
 DB CCCCAGCTGACGACCCCGCCAGCTGAATATGTACACAGCAGCATGTTCCGCCCCAGCCC 3447  
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
 DB CCTTCGCCCGGAGAGGGCCCTCTGCTGCTGCTGCCGACCTGCTGGTGCACCTCTGGAAGG 3507  
 QY 181 ProlysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
 DB CCAAGACTCTCTCCCGAGGAGAGATGGGGTGGTCAAAAGACGTTTTTGTGCTTTGGGGT 3567

QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220  
 DB GCCGTGGAGAACCCGAGTACTTGACACCCAGGAGGAGCTGCCCTCAGCCCCACCT 3627  
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240  
 DB CCTCCTGCTTCAGCCCGAGCTTCGACAACCTCTATTACTGGGACAGGACCCACAGAG 3687  
 QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
 DB CGGGGGGCTCCACCCAGCACCTTCAAAAGGGACACCTACGCGAGAGAACCCAGAGTACCTG 3747  
 QY 261 GlyLeuAspValProVal 266  
 DB GGTCTGGACGTGCCAGTG 3765  
 DB  
 RESULT 11  
 ID AAA14812  
 XX AAA14812 standard; cDNA; 4472 BP.  
 AC AAA14812;  
 XX 08-AUG-2000 (first entry)  
 DE cDNA encoding the SPLICE erbb-2 receptor protein.  
 XX  
 KW SPLICE erbb-2 receptor protein; cell transformation disorder; cancer;  
 KW tumor cell proliferation; tissue degeneration; arthropathy;  
 KW bone resorption; inflammatory disease; degenerative disorder;  
 KW wound healing; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 175..3942  
 FT /\*tag= a  
 FT /product= "SPLICE erbb-2 protein"  
 PN WO200020579-A1.  
 XX  
 PD 13-APR-2000.  
 XX  
 PF 01-OCT-1999; 99WO-CA00912.  
 XX  
 PR 02-OCT-1998; 98US-0165192.  
 XX  
 PA (UYMC-) UNIV MCMASTER.  
 XX  
 PI Muller WJ, Siegel PM;  
 XX  
 DR WPI; 2000-303768/26.  
 DR P-PSDB; AAY84780.  
 XX  
 XX Nucleic acid encoding an erbb 2 receptor protein designated SPLICE  
 PT erbb-2, inhibitors of the protein are useful for treatment of cancer -  
 XX  
 PS Claim 4; Fig 1; 60pp; English.  
 XX  
 CC The present sequence encodes a SPLICE erbb-2 receptor protein. The  
 CC protein has an in-frame deletion of 16 amino acids, 2 of which are  
 CC conserved cysteine residues, compared to the unspliced protein. The  
 CC erbb-2 polynucleotide is used to construct probes for detecting  
 CC disorders of cell transformation such as cancer. Antibodies to the  
 CC protein may be used to detect SPLICE erbb-2 in a sample. Agents  
 CC (e.g. antisense oligonucleotides) which inhibit the expression of  
 CC SPLICE erbb-2 are useful for reducing tumor cell proliferation and  
 CC treating cancer. Substances which stimulate SPLICE erbb-2 are useful  
 CC for treating conditions involving damaged cells including conditions  
 CC in which degeneration of tissue occurs, such as arthropathy, bone  
 CC resorption, inflammatory diseases, degenerative disorders of the  
 CC central nervous system and wound healing.  
 XX  
 SQ Sequence 4472 BP; 902 A; 1383 C; 1328 G; 859 T; 0 other;

DE	XX	Human tumour antigen ErBB-2 DNA.
KW	XX	Tumour antigen; human; vaccine; cellular immune response; immunogen;
KW	XX	cancer; tumour; ErBB-2; ds.
OS	XX	Homo sapiens.
PN	XX	US6287569-B1.
PD	XX	11-SEP-2001.
PF	XX	06-APR-1998; 98US-0056105.
PR	XX	10-APR-1997; 97US-043467P.
PA	XX	(REGC ) UNIV CALIFORNIA.
PI	XX	Kipps TJ, Wu Y;
DR	XX	WPI; 1998-583198/49.
PT	XX	Generating cellular immune response in patient to target protein -
PT	XX	comprises introducing vector with nucleotide sequence encoding
PT	XX	immunogen comprising protein processing signal into cell of patient
PS	XX	Disclosure; Column 81-86; 61pp; English.
CC	XX	This invention describes a novel method for generating a cellular immune
CC	XX	response in a patient to a target protein or its fragment. The method
CC	XX	involves introducing a vector containing a nucleotide sequence encoding
CC	XX	a chimeric immunogen comprising a protein processing signal and the
CC	XX	target protein or its fragment. The immunogen is produced by the cells
CC	XX	and processed so that the target protein or its fragment is presented to
CC	XX	the patients immune system and a cellular immune response is initiated.
CC	XX	The method and vectors can be used as a form of vaccination and could be
CC	XX	used to generate a cellular immune response in patients to, e.g.
CC	XX	cancerous tumours. The cellular immune response is the predominant immune
CC	XX	response in the patient. This sequence represents a DNA fragment which
CC	XX	encodes the human tumour antigen ErBB-2 described in the method of the
CC	XX	invention.
CC	XX	Note: The information in this spec has been previously disclosed in
CC	XX	WO199845444 however this spec contained no sequence information.
SQ	XX	Sequence 4473 BP; 902 A; 1383 C; 1329 G; 859 T; 0 other;
Alignment Scores:		
Pred. No.:	1.07e-66	Length: 4473
Score:	1450.00	Matches: 266
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	19	Gaps: 0
SEQ4 (1-266) x ABQ76220 (1-4473)		
QY	1	GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db	3142	CAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
QY	21	GluAspAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db	3202	GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 3261
QY	41	PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db	3262	TTCTTCTCTCCAGACCTGCCCGGGCGCTGGGGCATGGTCCACACAGCCAGCCGACG 3321
QY	61	SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db	3322	TCATCTACGAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3381
QY	81	AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db	3382	GGCCCGGAGTCTCCACTGGCACCTCCGAGGGGCTGGCTCCGATGATTTGATGGTGAC 3441
QY	101	LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db	3442	CTGGGAATGGGGCGGACCGAAGGGGTGCAAGCCCTCCACACATGACCCAGCCCTCTA 3501
QY	121	GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db	3502	CAGCGGTACAGTAGGAGCCACACAGTACCCCTGCGCTCTGAGACTGTGGTACGTTCGC 3561
QY	141	ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db	3562	CCCTTGACCTGCAGCCCGCCAGCCTGAATATGTGAACACAGCCAGATGTTCCGCCAGCC 3621
QY	161	ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db	3622	CCTTGGCCCGGAGAGGGCCCTCTGCTGCTGCCCGACCTGTGTGGTGCACCTCTGGAAGG 3681
QY	181	ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db	3682	CCCAAGACTCTCTCCCGAGGAAGAAATGGGTGCTGCAAGACGTTTTTTCCTTTGGGGT 3741
QY	201	AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db	3742	GCGGTGGAGAACCCCGAGTACTTGACACCCCGAGGGAGGAGTGCCCTCAGCCCACTCT 3801
QY	221	ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db	3802	CCTCTGCTTCCAGCCAGCCTTCGACAACTCTATTACTGGGACGAGGCCACCCAGAG 3861
QY	241	ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db	3862	CGGGGGGCTCCACCCAGCACCTTCAAAGGGACACCTACGGCAGAGAACCCAGAGTACCTG 3921
QY	261	GlyLeuAspValProVal 266
Db	3922	GCTCTGGACGTGCCAGTG 3939
RESULT 12		
ABQ76220		
ID	ABQ76220 standard; DNA; 4473 BP.	
XX		
AC	ABQ76220;	
XX		

21-OCT-2002 (first entry)  
Human tumour antigen ErbB-2 DNA.  
Tumour antigen; human; vaccine; cellular immune response; immunogen;  
cancer; tumour; ErbB-2; ds.  
Homo sapiens.  
OS  
US6287569-B1.  
11-SEP-2001.  
06-APR-1998; 98US-0056105.  
10-APR-1997; 97US-043467P.  
(REGC ) UNIV CALIFORNIA.  
Kipps TJ, Wu Y;  
WPI; 1998-583198/49.  
Generating cellular immune response in patient to target protein -  
comprises introducing vector with nucleotide sequence encoding  
immunogen comprising protein processing signal into cell of patient  
Disclosure; Column 81-86; 61pp; English.  
This invention describes a novel method for generating a cellular immune  
response in a patient to a target protein or its fragment. The method  
involves introducing a vector containing a nucleotide sequence encoding  
a chimeric immunogen comprising a protein processing signal and the  
target protein or its fragment. The immunogen is produced by the cells  
and processed so that the target protein or its fragment is presented to  
the patients immune system and a cellular immune response is initiated.  
The method and vectors can be used as a form of vaccination and could be  
used to generate a cellular immune response in patients to, e.g.  
cancerous tumours. The cellular immune response is the predominant immune  
response in the patient. This sequence represents a DNA fragment which  
encodes the human tumour antigen ErbB-2 described in the method of the  
invention.  
Note: The information in this spec has been previously disclosed in  
WO199845444 however this spec contained no sequence information.  
XX  
SQ Sequence 4473 BP; 902 A; 1383 C; 1329 G; 859 T; 0 other;

Alignment Scores:  
Pred. No.: 1.07e-66 Length: 4473  
Score: 1450.00 Matches: 266  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 19 Gaps: 0

SEQ4 (1-266) x ABQ76220 (1-4473)  
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
DB 3142 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGTCACTGCTG 3201  
QY 21 GluAspAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40  
DB 3202 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGGC 3261  
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60  
DB 3262 TTCTTCTCTCCAGACCTTCGCCCGGGCGCTGGGGCATGGTCCACACAGCCAGCCGACG 3321  
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
DB 3322 TCATCTACGAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3381



Db 3802 CCTCGTTCAGCCAGCCTTGCAGAACCTCTATTACTGGGACGAGCCACCCAGAG 3861  
 Qy 241 ArgGlyAlaProSerThrPheIysGlyThrProThrAlaGluAsnProGluTyrIleu 260  
 Db 3862 CGGGGGGCTCCACCCAGCACCTTCAAGGGACACCTAGCGGACAGAACCCAGAGTACCTG 3921  
 Qy 261 GlyLeuAspValProVal 266  
 Db 3922 GGTCTGGACGTGCCAGTG 3939  
 RESULT 14  
 ABZ34969  
 ID ABZ34969 standard; cDNA; 4473 BP.  
 AC ABZ34969;  
 DT 05-FEB-2003 (first entry)  
 DE Human gene expression profile polynucleotide SEQ ID NO 81.  
 KW Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;  
 KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;  
 KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;  
 KW gene expression; gene; ss.  
 XX Homo sapiens.  
 XX WO200274979-A2.  
 XX 26-SEP-2002.  
 XX 20-MAR-2002; 2002WO-US08456.  
 XX 20-MAR-2001; 2001US-276947P.  
 XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.  
 XX Wan J, Wang Y;  
 XX WPI: 2002-740862/80.  
 XX New gene expression profile generated from primary, endothelial,  
 XX epithelial, and muscle cell types, useful for identifying disease  
 XX pathologies involving alterations of gene expression, e.g. cancer  
 XX Claim 3; Page 274-276; 850pp; English.  
 XX The invention relates to a gene expression profile comprising one or more  
 XX genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type  
 XX is a coronary artery endothelium, umbilical artery or vein endothelium,  
 XX aortic endothelium, dermal microvascular endothelium, pulmonary artery  
 XX endothelium, myometrium microvascular endothelium, keratinocyte  
 XX epithelium, bronchial epithelium, mammary epithelium, prostate  
 XX epithelium, renal cortical epithelium, renal proximal tubule epithelium,  
 XX small airway epithelium, renal epithelium, umbilical artery smooth  
 XX muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,  
 XX dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,  
 XX aortic smooth muscle, mesangial cells, coronary artery smooth muscle,  
 XX bronchial smooth muscle, uterine smooth muscle, lung fibroblast,  
 XX osteoblasts or prostate stromal cell. The gene expression profile is used  
 XX for determining the level of RNA expression for a sample, determining the  
 XX phenotype of a cell and distinguishing cell types. The gene or a protein  
 XX expression profile is useful in identifying disease pathologies  
 XX involving alterations of gene expression. The assessment of expression  
 XX profiles may provide meaningful information with respect to tumour type  
 XX and stage, treatment methods, and prognosis. The gene or protein  
 XX expression profile may also be used for creating microarrays. The  
 XX microarray is useful for genetic and physical mapping of genomes, DNA  
 XX sequencing, genetic or medical diagnosis, genotyping of organisms,  
 XX confirming cell or tissue identifications and in identifying promising  
 XX antibiotics, antiviral or antifungal agents.

Seq Sequence 4473 BP; 902 A; 1383 C; 1329 G; 859 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1,07e-66 Length: 4473  
 Score: 1450.00 Matches: 266  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0  
 SEQ4 (1-266) x ABZ34969 (1-4473)  
 Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 Db 3142 CAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201  
 Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
 Db 3202 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3261  
 Qy 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60  
 Db 3262 TTCCTTCGTCGACAGCCCTCGCCCGGGGCTGGGGCATGGTCCACACAGGACCCGACG 3321  
 Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerSerGluGlu 80  
 Db 3322 TCATCTACCAGGAGTGGGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3381  
 Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
 Db 3382 GCCCCAGGCTCCACTGGCACCTCCGAAGGGCTGGCTCGCATGTATTTGATGGTGAC 3441  
 Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
 Db 3442 CTGGAAATGGGGGAGCCCAAGGGCTGCAAGGCTCCCAACATGACATGACCCAGCCCTCTA 3501  
 Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
 Db 3502 CAGCGGTACAGTGAGGACCCACAGTACCCCTGCTGAGACTGATGGCTACCTTGC 3561  
 Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
 Db 3562 CCCCTGACCTGACGCCCCAGCCTGAATATGTGAACACGAGAGATGTTGCGCCCGCAGCC 3621  
 Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
 Db 3622 CCTTGGCCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGGCCACTCTGGAAGG 3681  
 Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
 Db 3682 CCCAAGACTCTCTCCCGAGGAGAAATGGGGTGGTCTGCTCAAGAGAGTCTTTTTCCTTTGGGGGT 3741  
 Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220  
 Db 3742 GCCCTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGGTGCGCCCTCAGCCCGCCCT 3801  
 Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240  
 Db 3802 CCTCTGCTTTCAGCCAGCCCTTCGACAACTCTATTACTGGGACGAGCCAGGACCCAGAG 3861  
 Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
 Db 3862 CGGGGGGCTCCACCCAGCACCTTCAAGGGACACCTTCAAGGAGGACACCTACGGCAGAGAACCCAGAGTACCTG 3921  
 Qy 261 GlyLeuAspValProVal 266  
 Db 3922 GGTCTGGACGTGCCAGTG 3939  
 RESULT 15  
 AAD38904  
 ID AAD38904 standard; DNA; 4473 BP.  
 XX  
 AC AAD38904;

```

XX 23-SEP-2002 (first entry)
XX Human Her-2 DNA.
DE
DE
DE
KW Human; Her-2; epidermal growth factor receptor 2; infection; cancer;
KW hyperproliferative disorder; prophylaxis; inflammation; antisense;
KW tumour; gene therapy; phosphorothioate backbone; gene; ss.
XX
XX OS Homo sapiens.
XX
XX PH Key Location/Qualifiers
XX CDS 175..3942
XX FT /*tag= a
XX FT /product= "Human Her-2 protein"
XX
XX W0200222636-A1.
XX
XX PN
XX
XX PD 21-MAR-2002.
XX
XX PF 12-SEP-2001; 2001WO-US28572.
XX
XX PR 15-SEP-2000; 2000US-0663834.
XX
XX PA (ISIS-) ISIS PHARM INC.
XX
XX PI Bennett CF, Cowser LM;
XX
XX WPI: 2002-471192/50.
XX P-PSDB: AAE24067.
XX
XX Novel antisense oligonucleotide which modulates the expression of Human
XX Epidermal Growth Factor receptor, Her2, is useful for treating tumors
XX inflammation or to prevent infection in humans .
XX
XX Example 13; Page 94-101; 116pp; English.
XX
XX The invention relates to antisense compounds targetted to a nucleic
XX acid molecule encoding Her2 (human Epidermal Growth Factor receptor 2)
XX that specifically hybridises with and inhibits the expression of Her2.
XX Antisense compounds of the invention are used for treating diseases or
XX conditions associated with Her2 such as hyperproliferative disorders
XX e.g. lung, breast, gastric, oesophageal, colon, bladder, salivary,
XX neural or cardiac cancer. They are also useful prophylactically e.g.
XX to prevent or delay infection, inflammation and tumour formation. The
XX invention is also used in gene therapy. The present sequence is human
XX Her-2 DNA.
XX
XX SQ Sequence 4473 BP; 902 A; 1383 C; 1329 G; 859 T; 0 other.
XX
XX Alignment Scores:
XX Pred. No.: 1.07e-66 Length: 4473
XX Score: 1450.00 Matches: 266
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 24 Gaps: 0
XX
XX SEQ4 (1-266) x AAD38904 (1-4473)
XX
XX QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
XX Db 3142 CAGAAATGAGGACTGGGCCCGAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3201
XX
XX QY 21 GluAspAspMetGlyAspLeuValAspAlaGluClnTyrLeuValProGlnGlnGly 40
XX Db 3202 GAGACCATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCCGAGCGGC 3261
XX
XX QY 41 PhePheCysProAspProAlaProGlyAlaGlyClnMetValHisHisArgHisArgSer 60
XX Db 3262 TTCTTCGTCCAGACCTGCCCGGGCGCTGGGGCATGTGTCACACAGGACCGCAGC 3321
XX
XX QY 61 SerSerThrArgSerGlyGlyGlyAspLeuThrLeuGluProSerGluGlu 80

```

```

Db 3322 TCATCTACAGGAGTGGCGGTGGGGACCTGACACTAGGCTGGAGCCCTCTGAAGAGGAG 3381
QY 81 AlaProArgSerProLeuAlaProSerGluGlyValaGlySerAspValPheAspGlyAsp 100
Db 3382 GCCCCCAGGTCTCCACTGGCACCCTCCGAAAGGGCTGGCTCCGATGATTGATGGTGCAC 3441
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3442 CTGGGAATGGGGCAGCCAAAGGGCTGCAAAAGCTCCCCACACATGAGCCCGCCCTCTA 3501
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3502 CAGCGGTACAGTGGAGGACCCACAGTACCCCTGCGCTCTGAGACTGATGGCTAGCTTGCC 3561
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3562 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACAGGACAGATGTTCGGCCCCAGCCC 3621
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3622 CCTTCGCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGCCTCTGGAAGG 3681
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3682 CCCAAGACTCTCTCCCGAGGGAAGATGGGGTCTCAAGACGTTTTTGGCTTTGGGGGT 3741
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaPrcGlnProHisPro 220
Db 3742 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGTGCCTCCCTCAGCCCCACCCT 3801
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 3802 CCTCTGCTTCAGCCCGAGCCTTCGACAACCTCTATTACTGGGACGAGGCCACCCAGAG 3861
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3862 CGGGGGGCTCCACCCAGCACCTTCAAGAGGACACCTACGGCAGAGAACCCAGAGTACCTG 3921
QY 261 GlyLeuAspValProVal 266
Db 3922 GGCTGGACGTGCCAGTG 3939

```

Search completed: October 15, 2003, 23:50:06  
Job time : 277.648 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: October 15, 2003, 21:09:13 ; Search time 3830.63 Seconds  
(without alignments)  
2840.777 Million cell updates/sec

Title: SEQ4  
Perfect score: 1450  
Sequence: 1 QNEDLGASPLDSTFYRSLL.....TFKGTPTAENPEYLGLDV 266

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2889711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n model -DEV=xlp  
-Q/cgn2\_1/USPTO.spool\_p/HOLLERANA80/runat\_15102003\_131912\_20536/app\_query.fasta\_1.4685  
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERANA80.ecgn\_1.1.22724.@runat\_15102003\_131912\_20536 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -IGAPOP=10 -IGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_rst:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_rst:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rsd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	1450	100.0	1755	6	AX384609 Sequence
2	1450	100.0	1767	6	AX384607 Sequence
3	1450	100.0	1773	6	AX384610 Sequence
4	1450	100.0	3768	6	AR034479 Sequence
5	1450	100.0	3768	6	AX201817 Sequence
6	1450	100.0	3768	6	AX380923 Sequence
7	1450	100.0	3768	6	AX384604 Sequence
8	1450	100.0	3768	6	AX465456 Sequence
9	1450	100.0	4473	6	AR080259 Sequence
10	1450	100.0	4473	6	AR167390 Sequence
11	1450	100.0	4473	9	HSRBR2R Sequence
12	1442	99.4	1806	6	AX384608 Sequence
13	1442	99.4	3678	6	AX505114 Sequence
14	1442	99.4	3768	6	AX060704 Sequence
15	1442	99.4	3768	6	AX467229 Sequence
16	1442	99.4	3768	6	AX481438 Sequence
17	1442	99.4	4530	6	AR202597 Sequence
18	1442	99.4	4530	6	AR283481 Sequence
19	1442	99.4	4530	6	AX282577 Sequence
20	1442	99.4	4530	6	AX587649 Sequence
21	1442	99.4	4530	6	AX644071 Sequence
22	1442	99.4	4530	6	BD005474 Cellular
23	1442	99.4	4530	6	I21124 Sequence 9
24	1442	99.4	4530	6	I59745 Sequence 9
25	1442	99.4	4530	9	HUMHER2A Sequence
26	1442	99.4	9274	6	AX060703 Sequence
27	1358	93.7	161815	9	AC079199 Homo sapi
28	1358	93.7	168585	9	AC040933 Homo sapi
29	1358	93.7	198008	2	AC142197 Homo sapi
30	1350	93.1	13450	9	AB096612 Homo sapi
31	1350	93.1	16572	9	AB096614 Homo sapi
32	1350	93.1	20271	9	AB096613 Homo sapi
33	1350	93.1	30837	9	AY208911 Homo sapi
34	1291.5	89.1	3780	4	AB008451 Canis fam
35	1225.5	84.5	4062	10	HAMNEU Mesocricetu
36	1208	83.3	3955	6	AX380924 Sequence
37	1208	83.3	3955	6	I21129 Sequence 14
38	1208	83.3	3955	6	I59750 Sequence 14
39	1208	83.3	3955	10	RNNEUR X03362 Rat mRNA fo
40	1208	83.3	4727	10	AY116182 Rattus no
41	1205	83.1	2083	10	BC027080 Mus muscu
42	1205	83.1	2763	6	AX380942 Sequence
43	1205	83.1	2781	6	AX380944 Sequence
44	1205	83.1	3771	6	AX189662 Sequence
45	1205	83.1	3771	6	AX380925 Sequence

ALIGNMENTS

RESULT 1

```

AX384609
LOCUS AX384609 1755 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 6 from Patent WO0214503.
ACCESSION AX384609
VERSION AX384609.1 GI:19577810
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Hand-Zimmermann,S., Cheever,M.A., Foy,T.M., Lodes,M.J., Kalos,M.D.,
McNeill,P.D. and Vedvick,T.S.
TITLE Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 6 21-FEB-2002;
KEYWORDS CORIXA CORPORATION (US)
FEATURES
source
location/Qualifiers
1..1755
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 376 a 517 c 531 g 331 t
ORIGIN
Alignment Scores:
Pred. No.: 7.4e-54 Length: 1755
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ4 (1-266) x AX384609 (1-1755)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 946 CAGAATGAGGACTTGGGGCCACCGAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1005
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
DB 1006 GAGCAGCATGACATGGGGACCTGGTGATGCTCAGGAGTATCTGTGTACCCACGAGGC 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
DB 1066 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCATGTGTCACACAGCAGCCGAGC 1125
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
DB 1126 TCATCTACAGGAGTGGGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGGAAGAGGAG 1185
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
DB 1186 GCGCCCGAGGCTCCACTGGACACCTCCGAGGGGCTGGCTCCGATGATTTGATGGTGAC 1245
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
DB 1246 CTGGGAATGGGGGAGCAAGGGGCTGCAAGGCTCCACACATGACCCAGCCCTCTA 1305
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
DB 1306 CAGCGGTACAGTAGGACCCACAGTACCCCTGCCCTCTGAGACTGATGGCTACGTTGCC 1365
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
DB 1366 CCCCTGACCTGCACCCCGAGCTGATATGTGACAGCCAGATGTTGCGCCCGACGCC 1425
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
DB 1426 CCTTCGCCCCGAGAGGGCCCTCTGCCCTGCTGCCGACCTGTGTGTGCTCACTCTGGAAGG 1485
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
DB 1811 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly

```

```

DB 1486 CCCAAGACTCTCTCCCGAGGAGAAATGGGTCTCAAAGACGTTTTTGGCTTTGGGGGT 1545
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
DB 1546 GCCGTGGAGAACCCCGAGTACTTGCACACCCAGGAGAGCTGCTCCCTTCACCCCACT 1605
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
DB 1606 CCTCTGCCCTCAGCCCGAGCTTTCGACAACTCTATTACTGGGACGAGGCCACAGAG 1665
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
DB 1666 CGGGGGCTCCACCGACGACCTTCAAAGGGACACTACGGCAGAGAACCCAGAGTAGCTG 1725
QY 261 GlyLeuAspValProVal 266
DB 1726 GGTCTGGACGTGGCCAGTG 1743
RESULT 2
AX384607
LOCUS AX384607 1767 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 4 from Patent WO0214503.
ACCESSION AX384607
VERSION AX384607.1 GI:19577808
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Hand-Zimmermann,S., Cheever,M.A., Foy,T.M., Lodes,M.J., Kalos,M.D.,
McNeill,P.D. and Vedvick,T.S.
TITLE Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 4 21-FEB-2002;
FEATURES
source
location/Qualifiers
1..1767
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 381 a 521 c 529 g 336 t
ORIGIN
Alignment Scores:
Pred. No.: 7.45e-54 Length: 1767
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ4 (1-266) x AX384607 (1-1767)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 946 CAGAATGAGGACTTGGGGCCACCGAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 1005
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
DB 1006 GAGCAGCATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGTGTACCCACGAGGC 1065
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
DB 1066 TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCATGTGTCACACAGCAGCCGAGC 1125
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
DB 1126 TCATCTACAGGAGTGGGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGGAAGAGGAG 1185
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
DB 1186 GCGCCCGAGGCTCCACTGGCAGCCCTCCGAGGGGCTGGCTCCGATGATTTGATGGTGAC 1245

```

QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
 Db 1246 CTGGGAATGGGGGAGCCAGGCGTCAAGAGCTCCACACATGACCCAGCCCTCTA 1305  
 QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
 Db 1306 CAGCGGTACAGTGGAGGACCCACACAGTACCCCTGCTGAGACTGATGGCTAGCTTGC 1365  
 QY 141 ProLeuThrCysSerProGlnProGluTyrValAlaGlnProAspValArgProGlnPro 160  
 Db 1366 CCCCTGACTGACGCCCCCAGCCCTGAATATGTGAACAGCCAGATGTTGGCCCCAGCC 1425  
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
 Db 1426 CCTTCGCCCGAGAGGCGCTCTGCTGCTGCCGACCTGCTGCTGCTGCTGCTGCTGCTG 1485  
 QY 181 ProlysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
 Db 1486 CCCAAGACTCTCTCCCGAGGAGAAATGGGGTCTGCTCAAGAGCTTTTGGCTTTGGGGGT 1545  
 QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220  
 Db 1546 GCCGTGGAGAACCCGAGTACTTGACACCCAGGAGGAGCTGCCCTCAGCCACCCCT 1605  
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240  
 Db 1606 CCTCTGCTTCAGCCAGCTTCGACAACTCTATTACTGGGACGAGCCACCCAGAG 1665  
 QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
 Db 1666 CGGGGGGTCCACCCAGCACCTTCAAGGGGACCTACGGGAGAGAACCCAGAGTACCTG 1725  
 QY 261 GlyLeuAspValProVal 266  
 Db 1726 GGTCTGGACGTGCCAGTG 1743

RESULT 3  
 AX384610  
 LOCUS AX384610 1773 bp DNA linear PAT 19-MAR-2002  
 DEFINITION Sequence 7 from Patent WO0214503.  
 ACCESSION AX384610  
 VERSION AX384610.1 GI:19577811  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS Hand-Zimmermann,S., Cheever,M.A., Foy,T.M., Lodes,M.J., Kalos,M.D.,  
 McNeill,P.D. and Vedvick,T.S.  
 TITLE Compositions and methods for the therapy and diagnosis of  
 JOURNAL her-2/neu-associated malignancies  
 PATENT: WO 0214503-A 7 21-FEB-2002;  
 CORIXA CORPORATION (US)  
 FEATURES  
 Location/Qualifiers  
 source  
 1. 1773  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 383 a 528 c 530 g 332 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 7.47e-54 Length: 1773  
 Score: 1450.00 Matches: 266  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

SEQ4 (1-266) x AX384610 (1-1773)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 Db 970 CAGAAATGAGGACTTGGGCCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTACTGCTG 1029  
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40  
 Db 1030 GAGACCATGACATGGGGGACCTGGTGGATCTGAGGAGTATCTGGTACCCAGCAGGCG 1089  
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60  
 Db 1090 TTCTTCTGTCTCAGACCTTGCCTCCCGGGCTGGGGCATGTCTCCACACAGCAGCCGAC 1149  
 QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
 Db 1150 TCATCTACAGAGTGGCGGTGGGACCTGACACTAGGCTGGAGCCCTCTGAAGAGGAG 1209  
 QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAsnGlyAsp 100  
 Db 1210 GCCCCCAAGTCTCCACTGGCACCTTCCGAAGGGGCTGGCTCCGATGATTTGATGGTGAC 1269  
 QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120  
 Db 1270 CTGGAAATGGGGGAGCCAGGCGTCAAGAGCTCCACACATGACCCAGCCCTCTA 1329  
 QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
 Db 1330 CAGCGGTACAGTGGAGGACCCACAGTACCCCTGCTGAGACTGATGGTACGTTGCC 1389  
 QY 141 ProluthrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
 Db 1390 CCCCTGACTCTCCCGAGGAGAAATGGGGTCTGCTCAAGAGACGTTTTTGCCTTTGGGGT 1569  
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
 Db 1450 CCTTCGCCCGAGAGGCGCTCTGCTGCTGCCGACCTGCTGGTGGCACCTCTGGAAAGG 1509  
 QY 181 ProlysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
 Db 1510 CCCAAGACTCTCTCCCGAGGAGAAATGGGGTCTGCTCAAGAGACGTTTTTGCCTTTGGGGT 1569  
 QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220  
 Db 1570 GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCCCTCAGCCACCCCT 1629  
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240  
 Db 1630 CCTCTGCTCTCAGCCAGCCCTTGCACACCTCTATTACTGGGACGAGCCAGCCAGAG 1689  
 QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
 Db 1690 CGGGGGGCTCCACCCAGCACCTTCAAGGGGACCTACGGCAGAGAACCCAGAGTACCTG 1749  
 QY 261 GlyLeuAspValProVal 266  
 Db 1750 GGTCTGGACGTGCCAGTG 1767

RESULT 4  
 AR034479  
 LOCUS AR034479 3768 bp DNA linear PAT 29-SEP-1999  
 DEFINITION Sequence 1 from patent US 5869445.  
 ACCESSION AR034479  
 VERSION AR034479.1 GI:5950084  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 Unclassified.  
 REFERENCE 1 (bases 1 to 3768)  
 AUTHORS Cheever,M.A. and Disis,M.L.  
 TITLE Methods for eliciting or enhancing reactivity to HER-2/neu protein  
 JOURNAL Patent: US 5869445-A 1 09-FEB-1999;  
 FEATURES Location/Qualifiers  
 source  
 1. 3768  
 /organism="unknown"

```

BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN
Alignment Scores:
Pred. No.: 1.42e-53 Length: 3768
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

SEQ4 (1-266) x AR034479 (1-3768)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGCCCGGCGCTGGAGGACCTTACCCGCTCAGTCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGly 40
Db 3028 GAGGACGATGACATGGGGACCTGGTGATCTGAGGATATCTGGTACCCAGCAGGC 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db 3088 TTCTTCTGTCACAGACCTGCCCGGGCGCTGGGGCATGTCCACACAGCCAGCCGAGC 3147
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 80
Db 3148 TCATCTACGAGGATGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3207
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 100
Db 3208 GCGCCCGAGGCTCCACATGGCCCTCGAAGGGGCTGGCTCCGATGATTTGATGTGAC 3267
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuLysSerLeuProThrHisAspProSerPro 120
Db 3268 CTGGGAATGGGGCAGCAGGAGGCTGCAAGGCTCCCAACATGATGATGATGATGAT 3327
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrVal 140
Db 3328 CAGCGGTACAGTGGAGCCACACAGTACCCCTGCTGAGACTGATGCTACGTTGCC 3387
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGln 160
Db 3388 CCCTTGACCTGACGCCCCCAGCTGAATATGTGAACAGCAGCATGTTGCGCCCGAGCC 3447
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu 180
Db 3448 CCTTCGCCCCGAGAGGCCCTCTGCTGCTGCCGAGCTGCTGCTGCTGCTGCTGCTG 3507
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 200
Db 3508 CCCAAGACTCTCTCCCGAGGAGAAATGGGCTGCTCAAGACGCTTTTGGCTTTGGGGT 3567
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3568 GCGGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCCCTTCAGCCGCCCT 3627
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProPro 240
Db 3628 CTCTCTGCTTCAGCCAGCTTCGACAACTCTATTACTGGGACAGGACCCAGCAGAG 3687
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyr 260
Db 3688 CGGGGGGCTCCACCCAGCCTTCAAAGGAGACCTAGGCGAGAGAACCCAGAGTACCTG 3747
Qy 261 GlyLeuAspValProVal 266
Db 3748 GGTCTGGACGTGCAGTG 3765

RESULT 5
AX201817 3768 bp DNA linear PAT 30-AUG-2001
LOCUS
DEFINITION Sequence 1 from Patent WO0153463.

```

## ACCESSION

AX201817

## VERSION

AX201817.1 GI:15391666

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## source

## CDS

## codon\_start=1

## protein\_id="CAC60126.1"

## db\_xref="GI:15391667"

## translation="MELAAALCRWGLLLALLPGLAASQVCTGTDMKLRPLASPETHLD

## MLRLYGGQVQVQGNLELYTLPTNASLFDIQEVQGYVIIAHNQVRLPQRLRLTV

## RGQLFEDNYVALVLDNGDPLNNTPTVTGASPGGLRELQSLTEILKGGVLQRLRV

## LCYQDTILMKDIFPKNNQALATLIDTNRSPACHPCSPMKGSRGWSHSDCQSLRT

## VCAGGACRCKGLPTDCCHEOCAGCTGPKHSDCLACLHFNHSGICELHCPALVTNT

## DFESMPNPEGRTYFGASCTVACPYNTLSDVGSCTLVCPHMQEVTAEQGTQRCCK

## SKPCARVCGLGMHLEVRVAVTSANIQEPAGCKIFGSLAFIPESFDGDPASNTAPL

## QPQLQVFTLEITGYLXISAWPDSLPDLVSFQNLQVIRGRILHNGAYSLATQGLGI

## SWLGLSLRELGLALIHNTHLFCVHTVPMQDLFRNPHQALLHNRPEDECVGEG

## LACHOLCARGHCHGPGPTQCVNCSQFLRGQECVEECRVLQGLPREYVNAHCLPCHPE

## COPQNSVTCGCEPEADQVACAHVKDPPFCVAPRCSGVKPDLSYMPVLPKPEDEGACQ

## PCPINTHSCVDLDDKGCAPAEQASPLTSIISAVGILLVVLGVVFGILIKRQOKI

## RKYTMRLRLQETELVELPPLPSGAMPNOAMRIKTELKRVKVLGSGAFVTVKGIWI

## PDGENKIVAIKRLNTSPKANKEILDEAYVMAGVSPVSRILGICLCTSTVQLVT

## QLMPYCGCLLDHVRNDRGLSQDLNWCMIKAGMSYLEDVRLVLRDLARNVLSNP

## NHVKITDFGLARLLDIDETEHADGGKVPKIMWALESLIRRRDTHOSDVMSYGVYWE

## LMTFGAKPDVGIPARETDLLEKGERLPPOPICTIDVYIMVWKWIDVDSRFRREL

## VSFESRMARDPQREVIQNLDELGPASPLDSTFYRSILLEDDMDGLVDAEYGLVPQGF

## FCDDPAPAGGWHRRHSRSGGDLTLGLEPSEERAPRPLAPSEAGSDVFDG

## DLGMAAGLQSLPTDHPSPLOQYSEDTVPLPSETDGYVAPLTCSPQPEYVNOPDVR

## POPPSPREGPLPAARPAAGATLERPKTSLPGKNGVVKDFAFGGAVNPQEVITPQGGAA

## POPHPPATSPAFDNLVYWDQDPPERGAPSTFKGTPTAENPEYGLGLDVPV"

## BASE COUNT 759 a 1171 c 1119 g 719 t

## ORIGIN

## Alignment Scores:

## Pred. No.:

## Score:

## Percent Similarity:

## Best Local Similarity:

## Query Match:

## DB:

## SEQ4 (1-266) x AX201817 (1-3768)

## Qy

## Db

## Qy

## Db

## Qy

## Db

## Qy

## Db

## Qy

## Db

## Qy

## Db

## Qy

## Db

## Qy

## Db

## Qy

## Db

## Qy

## Db

## Qy

## Db

## Qy

## Db

## Qy

## Db

QY	81	AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp	100	
Db	3208	GCCCCAGGCTCCACCTGGCACCCCTCCGAAGGGCTGGCTCGATGATATTGATGGTGAC	3267	
QY	101	LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu	120	
Db	3268	CTGGGAATCGGGGACCCAGGGCTGCAGAGCTCCACACATGACCCCGCCCTCTA	3327	
QY	121	GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValala	140	
Db	3328	CAGCGGTACAGTGAGACCCACAGTACCCCTGCCTCTGAGACTGATGGCTAGTTGCC	3387	
QY	141	ProLeuThrCysSerProGlnProGluTyrValasnGlnProAspValArgProGlnPro	160	
Db	3388	CCCTGACCTGCAGCCCGCCAGCTGAATATGTGAACACGACGAGTGTTCGGCCCGCCG	3447	
QY	161	ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg	180	
Db	3448	CCTTCGCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGCCACTCTCGAAGG	3507	
QY	181	ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly	200	
Db	3508	CCCAAGACTCTCTCCCGAGGAGAAATGGGTGCTCAAGAGCTTTTGCTTTGGGGGT	3567	
QY	201	AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro	220	
Db	3568	GCGGTGGAGAACCCGAGTACTGTGACACCCAGGAGGAGCTGCCCTCGAGCCCGCCCT	3627	
QY	221	ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrPaspGlnAspProGlu	240	
Db	3628	CCTCCGCTCTCAGCCCGACCTTCGACAACTCTATTACTGGGACCCAGGACCCACGAG	3687	
QY	241	ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu	260	
Db	3688	CGGGGGCTCCACCCAGCACCTTCAAGGGGACACCTACGGCAGAGAACCCAGACTACTG	3747	
QY	261	GlyLeuAspValProVal	266	
Db	3748	GGTCTGGAGCTGCCAGTG	3765	
RESULT 6				
AX380923				
LOCUS	AX380923	3768 bp	DNA	linear
DEFINITION	Sequence 9 from Patent WO212341.			PAT 18-MAR-2002
ACCESSION	AX380923			
VERSION	AX380923.1	GI:19575767		
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1			
AUTHORS	Cheevar,M.A. and Gheysen,D.			
TITLE	Her-2/neu fusion proteins			
JOURNAL	Patent: WO 0212341-A 9 14-FEB-2002;			
FEATURES	CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)			
	Location/Qualifiers			
source	1..3768			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			
CDS	1..3768			
	/note="human Her-2/neu protein"			
	/codon_start=1			
	/protein_id="CAD28392.1"			
	/db_xref="GI:19575768"			
	/translation="MELAALCRWGLLALLPPGAASQVCTGTDMLKRLPASPEHLD			
	MURHLYQGVVQGNLELYLPTLQIEOVGYVLIAHNOVROVPLQRLRTV			
	RGTLFEENYALAVLDNGDLPLNTPTVPGASPGRLRELQLSLTEILKGVLTQRNP			
	LCYQDTILWKDIFHNQNALTLDTINRSRACHPCSPWCKSRGWSEKSDCSLRT			
	VGAGCARKGLPTDCCACGACGTPKRSDCGLACLFHNHSGICELHGPALVTYT			
	DTFESMPREGRYTFGASCVTACPNYLSLTDVSGCTLVCLPHNOEVTAEDTORCEK			

```

QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3508 CCCAAGACTCTCTCCCGAGGAGAAATGGGTGGTCAAGACGCTTTTGGCTTTGGGGT 3567
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3568 GCGGTGGAGAACCCGAGTACTTGACACCCAGGAGGAGCTGCCCTCAGCCGCCACCT 3627
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 3628 CCCTCCGCTTCAGCCAGGCTTCGACACCTCTATTACTGGACGAGGACCCACCCAGAG 3687
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3688 CGGGGGGCTCCACCCAGCCTTCAAGAGGACACCTACGCGAGAACCCAGAGTACCTG 3747
QY 261 GlyLeuAspValProVal 266
Db 3748 GGTCTGGACGTGCAGTG 3765

RESULT 7
AX384604
LOCUS AX384604 3768 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 1 from Patent WO0214503.
ACCESSION AX384604
VERSION AX384604.1 GI:19577806
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D., McNeill, P.D. and Vedvick, T.S.
TITLE Compositions and methods for the therapy and diagnosis of her-2/neu-associated malignancies
JOURNAL Patent: WO 0214503-A 1 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD28577.1"
/db_xref="GI:19577807"
/translation="MELALCRWGLLLALLPPGAASQTQVCTGTDMLRLPASPETHLD
MLRLYGGCVGVGNLELYLTPTNASLFLQDIOEVQVGLIAHNOVRQVPLQRLRV
RGTLPEADYVALVNDGDPNNTPVTGASPGRLRELQRLSLTEILKGVLLQRPQ
LCYQDTILMKDIEFKNNQALTLIDTNRSPACHPCPKWCSCWSESDEDCSLRT
VCAGCARCKGLPTDCCHQCAAGCTGPKHSDCLACLFHFNHSGICELHCPALVYNT
DTESMPNPGRTYVASCVTAQPNLYLTDVSGCTLVCPLNHQNVEADGTQRCCK
SKPCARVGLGMEHLREVRVNTSANTQEFAGCKIFGLAFELPESFDGDPASNTAPL
QPQLQVFTLEITGLYLSAMPDSDLPDLVFQNLQVIRLTHNGAYSLTLQGLI
SWGLSLRLGSGSLALIHNTHLCLFVHTVPDQLFNPQALHTANRDECEVGE
LACHOLCARGCHGPGPTQCVNCSQFLRGCEVECRVQLGLPREYVNRHCLPCHPE
CQNCNTHSCGDFEADQCVACAHKDPFCVAPRCPSGVKPDLSIMPIKFPDDEGACQ
PCPNTHSCDLDLDDCPAEQASPLTSIISAVGILLVILVGVFGLIKRRQKI
RKYTKRRLQETELVPLTPSGAMNQAKRIKTELKRVKLVGSYVSLGICLTSTVLT
PDGENTKRIKVAIKRENTSPKANKELDEAYVMAGVSPYVSRLLGICLTSTVLT
OLMPYCLDHLVRENKRLGSDLLNWCQIAKMSYLEDVRLVHRLDARNVLPKSP
NHVKITDGLARLLDIEYHADGKVPKIKWALSIILRRRTHOSDWSYGVTVWE
LMTFGAKPDGIFAREIPDLLEKGRIPQPICTIDYIMVWKWIMDSECRFRREL
VSFSRMARDPQRFVYQNEIDGLPGLDSTFYRSLLEDODMDGLVDAYEVLVPOQGF
FCPDPAFGAGVYHRRHSSTRSGGDLTLGLEPSEEAAPRSLAPSEGAGSDVDFG
DLGWAAGKQLSLTPHDSPQLQYSEDVPLPSETDGVVAPLTCSPOPEYVNOPVR
PQPSREGPLPAARAGATLPRPKTLLSPKNGVVKDVFAGFAGAVENPEYLTQGGAA
PQPPPPAESPADNLYWDDDPERGAPSTFKGTPTAENPEYGLDVPV"
59 a 1171 c 1119 g 719 t

```

```

Alignment Scores:
Pred. No.: 1 42e-53 Length: 3768
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

SEQ4 (1-266) x AX384604 (1-3768)
QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CACAATGAGGACTGGGCCAGCCAGTCCCTTGACAGCAGCACCTTCTACCGCTCAGTCTG 3027
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGC 3087
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60
Db 3088 TTCTTCTGTCACACCTGCCCGGGGCTGGGGGATGCTCCACACAGCAGCCAGCAG 3147
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3148 TCATCTACCAGGAGTGGCGGTGGGACCTGACACTAGGAGTGGAGGCTCTGAAGAGGAG 3207
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3208 GCCCCAGGTCCTCCTACCTGGCACCTCCGAAAGGGGCTGGCTCCGATGATTTGATGGT 3267
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3268 CTGGGAATGGGGGAGCCAGAGGGCTGCAAGGCTCCACACACATGACCCAGCCCTTA 3327
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3328 CAGCGGTACAGTGGAGGACCCAGTACCCCTGCCCTCTGAGACTGATGGCTACGTTGCC 3387
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3388 CCCCTGACCTGACGCCCCCAGCTGATATGTGAACAGCAGATGTTCCGCCGCCAGCCC 3447
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3448 CCTTGGCCCGCAGAGGGGCCCTCTGCCCTGCTGCCGACCTGCTGGTGCACCTCTGGA 3507
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3508 CCCAAGACTCTCTCCCGAGGAGAAATGGGTGGTCAAGACGCTTTTGGCTTTGGGGT 3567
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3568 GCGGTGGAGAACCCGAGTACTTGACACCCAGGAGGAGCTGCCCTCAGCCGCCACCT 3627
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 3628 CCCTCCGCTTCAGCCAGGCTTCGACACCTCTATTACTGGACGAGGACCCACCCAGAG 3687
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3688 CGGGGGGCTCCACCCAGCCTTCAAGAGGACACCTACGCGAGAACCCAGAGTACCTG 3747
QY 261 GlyLeuAspValProVal 266
Db 3748 GGTCTGGACGTGCAGTG 3765

RESULT 8
AX465456
LOCUS AX465456 3768 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 1 from Patent WO0213847.
ACCESSION AX465456
VERSION AX465456.1 GI:21899819

```



```

QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyLeuValProGlnGlnGly 40
Db 3202 GAGGACGACATGGGGGACCTGGTGATGCTGAGAGATATCTGGTACCCAGCAGGC 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgSer 60
Db 3262 TTCTTCTGTCAGACCTGCCCGGGCGCTGGGGCATGGTCCACACAGCCAGCCGAGC 3321
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 80
Db 3322 TCATCTACAGGAGTGGCGGTGGGACCTGACACTAGGGGTGGAGCCCTCTGAAGAGGAG 3381
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 100
Db 3382 GCCCAGAGTCTCCACTGGACCCCTCGAAGGCTGCCATGCCATGATTTGATGTGAC 3441
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 120
Db 3442 CTGGGAATGGGGCAGCAAGGGCTGCAAGGCTCCCAACATGACCCAGCCCTCTA 3501
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrVal 140
Db 3502 CAGCGGTACAGTGGAGCCCTGACACTGACCTGCTGCTGCTGCTGCTGCTGCTGCTG 3561
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGln 160
Db 3562 CCCTTGACCTGCAGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3621
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu 180
Db 3622 CCTTCGCGCCGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3681
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 200
Db 3682 CCCAAGACTCTCTCCCGAGGAGAGATGGGTGCTCAAGACGCTTTTGGCTTTGGGGT 3741
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3742 GCCGTGGAGAACCCGAGTACTTGCACACCTTATTTACTGGGACGAGCCACCCAGAG 3801
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
Db 3802 CTCTCTGCTTACGCCGAGCTTCGACAACTCTATTACTGGGACGAGCCACCCAGAG 3861
QY 241 ArgGlyAlaProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3862 CGGGGGCTCCACCCAGCCTTCAAAGGACACCTAGGCGAGAGAACCCAGAGTACCTG 3921
QY 261 GlyLeuAspValProVal 266
Db 3922 GGCTGGACGTGCCAGTG 3939

```

```

RESULT 10
AR167390 4473 bp DNA linear PAT 17-DEC-2001
LOCUS
DEFINITION Sequence 26 from patent US 6287569.
ACCESSION AR167390
VERSION AR167390.1 GI:17903168
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 4473)
AUTHORS Kippes,T.J. and Wu,Y.
TITLE Vaccines with enhanced intracellular processing
JOURNAL Patent: US 6287569-A 26 11-SEP-2001;
FEATURES
source Location/Qualifiers
1..4473
/organism="unknown"
BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN
Alignment Scores:

```

```

Pred. No.: 1.64e-53 Length: 4473
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SEQ4 (1-266) x AR167390 (1-4473)
QY 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGAATCAGGACTTGGGCCAGCCAGTCCCTTGACAGCACCTTCTTACCGCTCACTGCTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyLeuValProGlnGlnGly 40
Db 3202 GAGGACGATGACATGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGC 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgSer 60
Db 3262 TTCTTCTGTCAGACCTGCCCGGGCGCTGGGGCATGGTCCACACAGCCAGCCGAGC 3321
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 80
Db 3322 TCATCTACAGGAGTGGCGGTGGGACCTGACACTAGGGGTGGAGCCCTCTGAAGAGGAG 3381
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3382 GCCCAGAGTCTCCACTGGACCCCTCGAAGGCGCTGCCATGATTTGATGTGAC 3441
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 120
Db 3442 CTGGGAATGGGGCAGCAAGGGCTGCAAGGCTCCCAACATGACCCAGCCCTCTA 3501
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrVal 140
Db 3502 CAGCGGTACAGTGGAGCCCTGACACTGACCTGCTGCTGCTGCTGCTGCTGCTGCTG 3561
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGln 160
Db 3562 CCCTTGACCTGCAGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3621
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu 180
Db 3622 CCTTCGCGCCGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3681
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 200
Db 3682 CCCAAGACTCTCTCCCGAGGAGAGATGGGTGCTCAAGACGCTTTTGGCTTTGGGGT 3741
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3742 GCCGTGGAGAACCCGAGTACTTGCACACCTTATTTACTGGGACGAGCCACCCAGAG 3801
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
Db 3802 CTCTCTGCTTACGCCGAGCTTCGACAACTCTATTACTGGGACGAGCCACCCAGAG 3861
QY 241 ArgGlyAlaProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3862 CGGGGGCTCCACCCAGCCTTCAAAGGACACCTAGGCGAGAGAACCCAGAGTACCTG 3921
QY 261 GlyLeuAspValProVal 266
Db 3922 GGCTGGACGTGCCAGTG 3939

```

```

RESULT 11
HSEB2R
LOCUS
DEFINITION Human c-erb-B-2 mRNA.
ACCESSION X03363
VERSION X03363.1 GI:31197
KEYWORDS cell surface glycoprotein; cellular oncogene; erb-2 cellular; glycoprotein; growth factor receptor; neu cellular

```





```

Db      3862  CGGGGGCTCCACCAGCACCCTTCAAAGGGGACACCTACGGCAGAGAACCCAGAGTACCTG 3921
Qy      261  GlyLeuAspValProVal 266
          |||||||
Db      3922  GGTCTGGACGTGCCAGTG 3939

RESULT 12
AX384608
LOCUS      AX384608                1806 bp            DNA            linear      PAT 19-MAR-2002
DEFINITION Sequence 5 from Patent WO0214503.
ACCESSION  AX384608
VERSION     AX384608.1  GI:19577809
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     McNeill,P.D. and Vedvick,T.S.
TITLE       Hand-Zimmermann,S., Cheever,M.A., Foy,T.M., Lodes,M.J., Kalos,M.D.,
            Compositions and methods for the therapy and diagnosis of
            her-2/neu-associated malignancies
JOURNAL     Patent: WO 0214503-A 5 21-FEB-2002;
            CORIXA CORPORATION (US)
FEATURES    source
            location/Qualifiers
            1..1806
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT 391 a 530 c 544 g 341 t
ORIGIN

Alignment Scores:
Pred. No.:      1.66e-53      Length:      1806
Score:          1442.00      Matches:    265
Percent Similarity: 99.62%      Conservative: 0
Best Local Similarity: 99.62%      Mismatches: 1
Query Match:     99.45%      Indels:    0
DB:              6          Gaps:      0

SEQ4 (1-266) x AX384608 (1-1806)

Qy      1  GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
          |||||||
Db      1003  CAGAATGAGGACTTGGGCCAGCAGCTCCCTTGCAGCAGCACCTTCTACCGCTCACTGCTG 1062
Qy      21  GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
          |||||||
Db      1063  GAGGACGATCAGATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCCGCAGAGGC 1122
Qy      41  PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
          |||||||
Db      1123  TTCTTCGTCCAGACCTGCCCGCGGCTGGGGGATGTCACACAGGCACCGCAGC 1182
Qy      61  SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
          |||||||
Db      1183  TCATCTACCAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 1242
Qy      81  AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
          |||||||
Db      1243  GCCCCCAGGTCTCCATGGCACCCTCGAAGGGGCTGGCTCCGATGATTTGATGGTGAC 1302
Qy      101  LeuGlyMetGlyAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
          |||||||
Db      1303  CTGGGAATGGGGCAGCAAGGGCTGCAAAAGCCTCCCCACATGACCCCGCCCTCTA 1362
Qy      121  GluArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
          |||||||
Db      1363  CAGCGGTACAGTGAAGACCCACAGTACCCCTCCCTCTGAGACTGATGGCTACGTGCC 1422
Qy      141  ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
          |||||||
Db      1423  CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACAGCCAGCATGTTGGCCCCCAGCCC 1482

```

```

Qy      161  ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
          |||||||
Db      1483  CCTCGCCCGCAGAGGGCCCTCTGCTGCTGCTGCCGACCTGCTGCTGCCACTCTCGAAGG 1542
Qy      181  ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
          |||||||
Db      1543  GCCAAGACTCTCTCCCGAGGGAAGATGGGCTGCTCAAAGACGTTTTTGGCTTTGGGGGT 1602
Qy      201  AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
          |||||||
Db      1603  GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGCTGCCCTCTACGCCCCACCT 1662
Qy      221  ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
          |||||||
Db      1663  CTTCTCTGCTTCAGCCCGCCTTCGACAACTTATTACTGGGACAGGACCCACACAGAG 1722
Qy      241  ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
          |||||||
Db      1723  CGGGGGCTCCACCCAGCACCTTCAAAGGAGACCTACGGCAGAGAACCCAGAGTACCTG 1782
Qy      261  GlyLeuAspValProVal 266
          |||||||
Db      1783  GGTCTGGACGTGCCAGTG 1800

RESULT 13
AX505114
LOCUS      AX505114                3678 bp            DNA            linear      PAT 27-SEP-2002
DEFINITION Sequence 7 from Patent WO0240059.
ACCESSION  AX505114
VERSION     AX505114.1  GI:23386421
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Mincheff,M.S., Loukinov,D.I. and Zoubak,S.
TITLE       Methods. and compositions for inducing cell-mediated immune
            responses
JOURNAL     Patent: WO 0240059-A 7 23-MAY-2002;
            American Foundation for Biological Research Inc. (US) ; Mincheff,
            Milcho S. (US) ; Loukinov, Dmitri I. (US) ; Zoubak, Serguei (US)
FEATURES    source
            location/Qualifiers
            1..3678
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            7..3678
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD48804.1"
            /db_xref="GI:23386422"
            /translation="MRLPASPTHLDMLRHLRYOGCQVVGQNLTYLPTNWSLSFLQD
            IQEVQYVLIAHNQVRQVPLRIYRGVLTQLEDNVALAVLDNGDPLNNPTPYTGASP
            GGLRIQLRSLRLEILKGGVLIQRPOLQDTILWKDIFHKNQALTLTDTRSRAC
            HPCSPCKGSRGCGSESDCOSLTRVCAGCARCKGPIPLDCCHEOCAGCTGPKHS
            DCLACLFHNFHSGICELHCALVTYNTDTFESMPNBRGYTFGACVYACPYLSTDV
            GSCTLVCPLHNOEVTAEADGTORCEKCKPCARVYCGLGMHLEVRVAVTSANTQEFAG
            CKKIFGSLAPLPESTFGDPSANTAPLQPOLQVFETLEETITGLYLSAWPDSLPLSV
            FQNLQVIRKLHNGAYSUTLQGLGISWLGRSLRLEGLSGALIHHTHLCFVHTVPM
            DQLFRNPHALHTANRDECEGEGELACHQCHGWCVPQTQCVNCSQFRLGQEC
            VESCRVQLGLPREYNARHCLPCHPECPQNGSVTCFGEADQCVACAHKDPKPCVA
            RCPGKVPDLSPYMPWKFPDEGACOPCINCTHSCVDLDDKCGCPAORASPLTSIVS
            AVVGILLVVVLGVVFCGLIKRRQOKIRKVTMRLLQETELVEPLTSGAMPNQAOMRI
            LKTELRLKVKLGSGAFGVYKGINIPDGENVKIPVAIKVLRNTPSPKANKEILDENY
            VMAGVSPYVSRLLLGICLTSTVOLTPYGLDHLHVRNRLGSLQDLMNCMQIA
            KGMYSLEDVRLVRLHDLAARNLVKSPNHVKTIDFLARLIDIDITYHADGGVPIKW
            MALESILRRFTHQSDVWSYGVTVVWELMTFGAKPYDGIPIAREIPLDLEKERLPQPI
            CTIDVYMWVKWCMIDSECRPRELSEFSRWARDPQRFVIONLEDGLPASPLDSTF
            YRSLLEDDMDGLVDAEEVLPQOGFCFCDPDPAPCAGMVHHRHSSTRSGSGDGLTLG
            LEPSSEAPRSPPLAPSEGAGSDVFDGLMGAAKGLQSLPTHDSPLQRYRSDEPTVPL
            PSTETDGYVAPLTCSPQPEYVNPQDVRPQPPSPREGPLPAPARPAGATLERAKTILSPGKN

```

GVKDVFAFGAVENPEYLTTPQGAAPQHPHPPAFSPAFDNLYWDQDPPPERGAPPST  
FKGTPTEAPENPEYLGIDVPV"  
BASE COUNT 746 a 1137 c 1091 g 704 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.04e-53 Length: 3678  
Score: 1442.00 Matches: 265  
Percent Similarity: 99.62% Conservativeness: 0  
Best Local Similarity: 99.62% Mismatches: 1  
Query Match: 99.45% Indels: 0  
DB: 6 Gaps: 0

SEQ4 (1-266) x AX505114 (1-3678)

```
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2878 CAGAAATGAGGACTTGGGCCCGACCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 2937
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 2938 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCACGACGGGC 2997
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db 2998 TTCTTCTCTCCAGACCTTCCCGCGGCGCTGGGGCATGGTCCACACGACCGCGCAGC 3057
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGluGluProSerGluGluGlu 80
Db 3058 TCATCTACCGAGGTGGGGGACCTGACATAGGGCTGGAGCCCTCTGAAAGAGGAG 3117
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3118 GCGCCCGAGTCTCCACTGGCACCTTCCGAGGGGCTGGCTCCGATGATTTGATGGTGAC 3177
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3178 CTGGGAATGGGGGAGCGCAAGGGCTGCAAAAGCTCCACACATGACCCCGACCCCTCTA 3237
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3238 CAGCGGTACAGTGGAGGCCACACAGTACCCCTGCGCTCTGAGACTGATGGCTAGCTGGCC 3297
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3298 CCGCTGACCTGCGACCCCGACCCCTGAATATCTGAAACGACGATGTTGGCCCGCCAGCCC 3357
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3358 CCTTCGCGCCGAGAGGGGCTCTGCTGCTGCGCGACCTGCTGGTGGCCACTCTGGAAGG 3417
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3418 GCCAAGACTCTCTCCCGAGGAGAAATGGGGCTGCTCAAAAGACGTTTGGCTTTGGGGGT 3477
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3478 GCCGTGGAGAACCCGAGTACTTGACACCCCGAGGAGGAGTGGCCCTCACCCCGACCCCT 3537
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
Db 3538 CCTCCTGCTTTCAGCCCGACCTTTCGACAACTCTATTACTGGGACCGACCGACCGAG 3597
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3598 CGGGGGGCTCCACCGACACCTTCAAAGGGGACCTACGGGAGAGAACCCCGAGAGTACCTG 3657
Qy 261 GlyLeuAspValProVal 266
Db 3658 GGTCTGGACGTGCCAGTG 3675
```

RESULT 14  
AX060704

LOCUS AX060704 3768 bp DNA linear PAT 22-JAN-2001  
DEFINITION Sequence 2 from Patent WO0100244.  
ACCESSION AX060704  
VERSION AX060704.1 GI:12406101  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Erickson, S. and Schwall, R.  
TITLE Methods of treatment using anti-erbB antibody-maytansinoid  
conjugates  
JOURNAL Patent: WO 0100244-A 2 04-JAN-2001;  
Genentech, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..3768  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 758 a 1170 c 1121 g 719 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.1e-53 Length: 3768  
Score: 1442.00 Matches: 265  
Percent Similarity: 99.62% Conservativeness: 0  
Best Local Similarity: 99.62% Mismatches: 1  
Query Match: 99.45% Indels: 0  
DB: 6 Gaps: 0

SEQ4 (1-266) x AX060704 (1-3768)

```
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAAATGAGGACTTGGGCCCGACCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCACGACGGGC 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db 3088 TTCTTCTCTCCAGACCTTCCCGCGGCGCTGGGGCATGGTCCACACGACCGCGCAGC 3147
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGluGluProSerGluGluGlu 80
Db 3148 TCATCTACCGAGGTGGGGGACCTGACATAGGGCTGGAGCCCTCTGAAAGAGGAG 3207
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3208 GCGCCCGAGTCTCCACTGGCACCCCTCCGAGGGGCTGGCTCCGATGATTTGATGGTGAC 3267
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3268 CTGGGAATGGGGGAGCGCAAGGGCTGCAAAAGCTCCCGACACATGACCCCGACCCCTCTA 3327
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3328 CAGCGGTACAGTGGAGACCCCGACAGTACCCCTGCGCTCTGAGACTGATGGCTAGCTTGGCC 3387
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3388 CCGCTGACCTGACCGCCCGACCCCTGAATATGTGAACCGACAGATGTTGGCGCCCGACGCC 3447
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3448 CCTTCGCGCCGAGAGGGGCTCTGCCCTGCTGCCGACCTGCTGGTGGCGACCTCTCGAAGG 3507
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3508 GCCAAGACTCTCTCCCGAGGAGAAATGGGGTCTGCTCAAAGACGTTTTTGGCTTTGGGGGT 3567
```

QY 201 AlavalgluasnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220  
 Db 3568 GCGTGGAGAACCCCGAGTACTTGCACACCCAGGAGGAGCTGCCCTCAGCCCAACCT 3627  
 QY 221 ProProAlaPheSerProAlaPheAsnLeuTyrTrpAspGlnAspProProGlu 240  
 Db 3628 CTTCTGCGCTTCAGCCAGCCTTCGACAACTCTATTACTGGACAGGACCCACAGAG 3687  
 QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
 Db 3688 CGGGGGGCTCCACCCAGCCTTCAAGAGGACACCTAGGCGAGAACCCAGAGTACCTG 3747  
 QY 261 GlyLeuAspValProVal 266  
 Db 3748 GGTCTGGACGTGCAGTG 3765

RESULT 15  
 AX467229 3768 bp DNA linear PAT 16-JUL-2002  
 LOCUS  
 DEFINITION Sequence 3 from Patent WO0234287.  
 ACCESSION AX467229  
 VERSION AX467229.1 GI:21900511  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 1  
 Beier, A.M., Gautam, A. and Mouritsen, S.R.  
 TITLE Novel therapeutic vaccine formulations  
 JOURNAL Patent: WO 0234287-A 3 02-MAY-2002;  
 Pharmexa A/S (DK)

FEATURES  
 source  
 1..3768  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 CDS  
 1..3768  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD42452.1"  
 /db\_xref="GI:21900512"  
 /translation="MELALCRWGLLLALLPCCAASVCTGTDMKLRPLPASPETHLD  
 MRHLQGVQVQGNLELYLPNASLSLQDIOEVGVLIAHNOVQVPLQRRLV  
 RGQLFEDNTALAVLNDPLNNTPTVTSASGLRELQRLSLTEILKGGVLIQRNPQ  
 LCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPMCKSGRCWGESSEDCQSLVPT  
 VCAGGARGCKGPLPTDCEQAAGCTGPKHSDCLACLFHNSGICELHCPALVYNT  
 DTESMPNPEGRVTFGASCTVACPNVLTSDVGSCTLVGLPHNOEVTAEQGTQCEKC  
 SKPCARVCYGLGMEHLREVRVTSANIQEFAGCKIFGSLAPLPESFDGDPASNTAPL  
 QPOLQVFTLEITGLYXISAMPDSLPLSVQNLQVIRGRILLHNGAYSLSLQGLGI  
 SWLGLSLRELHGLALIHNTHLCEVHTVPMQDLFRNPQHLHNTANRDECEVGE  
 LACHQLCARGHCWPGPTQVCNCSQFLRQECVEECVQLGFLPREYVNAHRLCPHPE  
 COPQNSVTFCGPEADQVCAHYKDPFCVAPCSGVKPDLSYMPYIKWFPDEGACQ  
 PCPINTHSCVDLDDKGCFAEORASPLTSIVSAVGGILLVWLVVFGILIKRROKI  
 RKYTMRLQETELREPLTPSGAMPNOAMRILKETELBKVKVLSGAGVYKGIWI  
 PDGENVKIPVAKLVRENTSPRANKELDEAVYMGVSPVSRLLGICLTSTVQLVT  
 QLMPTGLDHLVRENRLGSLDNLNMQIAKMSYLEDLVLRHDSVMSYGVYWE  
 NHVYITDGLRLDIDETEYHADGKVPYIKWMALESILRRFTHQSDVMSYGVYWE  
 LMTFGAPKIDGIPAREPDLLEKGLPQPPICITIDVIMVIMKWMIDSECRPRREL  
 VSEFSRMAKQVRFVIONLEKGLPQPPICITIDVIMVIMKWMIDSECRPRREL  
 FCPDPAAGGAGVHHRSSSTFGGGDLTLGLPESEEPAPSPAPSGAGSDVFDG  
 DLGMAKAGLSLPTHTDPLQRYSEDVPLPSETDGVYVPLTCSQPEYVNPQDVR  
 PQPSPREGPLPAARPAALERAKTLSPGKNGVKVDVFAFGVAGNPEYLTLPQGGA  
 PQPHPPAPSPAFDNLVYWDQDPPGAPSPFTKGTPTAENPEYLGIDVYV"

BASE COUNT 758 a 1170 c 1121 g 719 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3,1e-53 Length: 3768  
 Score: 1442.00 Matches: 265  
 Percent Similarity: 99.62% Conservatives: 0  
 Best Local Similarity: 99.62% Mismatches: 1  
 Query Match: 99.45% Indels: 0

Search completed: October 16, 2003, 11:03:03  
 Job time: 3853.81 secs

DB: 6 Gaps: 0  
 SEQ4 (1-266) x AX467229 (1-3768)  
 QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20  
 Db 2968 CAGAAATGAGACTTGGGCCCCAGCCAGTCCCTTGGACAGCACCTTCTACCCGCTCACTGCTG 3027  
 QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40  
 Db 3028 GAGGACGATGACATGGGGACCTGGTGATGCTGAGAGTATCTGGTACCCAGCAGGGC 3087  
 QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60  
 Db 3088 TTCTTCTGTCCAGACCTGCCCCGGCGCTGGGGGATGTTGCCACACAGGACCCGAGC 3147  
 QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80  
 Db 3148 TCATCTACCAAGGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGGAGG 3207  
 QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100  
 Db 3208 GCCCCAGGCTCTCCACTGGCACCTCCGAAAGGGCTGGCTCCGATGTATTTGATGGTGAC 3267  
 QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisaspProSerProLeu 120  
 Db 3268 CTGGGAATGGGGGACGACCAAGGGGCTGCAAAAGGCTCCCCACACATGAGCCACGCCCTCTA 3327  
 QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140  
 Db 3328 CAGCGGTACAGTACGAGACCCACAGTACCCCTCTGAGACTGNTGGCTACGTGGCTGCC 3387  
 QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160  
 Db 3388 CCCCTGACCTGACGCCGCCAGCTGAATATGTCAACAGCAGATGTTCCGGCCCCAGCCC 3447  
 QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180  
 Db 3448 CCTTCGCCCGGAGAGGGGCCCTCTGCTGCTGCCGACCTGCTGCTGCCACTCTGGAAGG 3507  
 QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200  
 Db 3508 GCCAAGACTCTCTCCCCAGGGAAGATGGGCTGCTCAAGACCGTTTTTGGCTTTGGGGGT 3567  
 QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220  
 Db 3568 GCGGTGGAGAACCCCGAGTACTTGCACACCCAGGAGGAGCTGCCCTCAGCCCCACCT 3627  
 QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240  
 Db 3628 CTTCTGCGCTTCAGCCAGCCTTTCGACAACTCTATTTACTGGACAGGACCCAGGAC 3687  
 QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260  
 Db 3688 CGGGGGGCTCCACCCAGCAGCTTCAAGAGGACACCTACGGCAGAGAACCCAGAGTACCTG 3747  
 QY 261 GlyLeuAspValProVal 266  
 Db 3748 GGTCTGGACGTGCAGTG 3765

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 5110.68 Seconds  
(without alignments)  
3105.423 Million cell updates/sec

Title: SEQ3  
Perfect score: 3628  
Sequence: 1 MELAALCRWGLLLALLPPGA.....CVDLDDKCPAERASPLTS 653

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -MODEL=frame\_plus\_p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO\_spool\_P/HOLLERAN480/runat\_15102003\_131913\_20548/app\_query.fasta\_1.4685  
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 -CGN\_1\_1\_13645 -runat\_15102003\_131913\_20548 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_fod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gssl.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3087.5	85.1	4323	11 AK031099	AK031099 Mus muscu
2	3087.5	85.1	4463	11 AK083669	AK083669 Mus muscu
3	1533.5	42.3	2456	11 AK004911	AK004911 Mus muscu
4	1533.5	42.3	2662	11 AK004883	AK004883 Mus muscu
5	1533.5	42.3	2936	11 AK004944	AK004944 Mus muscu
6	1483	40.9	964	13 BUI150809	BUI150809 AGENCOURT
7	1410	38.9	1016	12 BM562913	BM562913 AGENCOURT
8	1364	37.6	757	9 AU140362	AU140362 AU140362
9	1339	36.9	3110	11 AK031542	AK031542 Mus muscu
10	1333	36.7	808	14 CA489534	CA489534 AGENCOURT
11	1293	35.6	894	10 BE746725	BE746725 601579159
12	1205.5	33.2	795	14 CA328613	CA328613 UI-M-FY0-
13	1194.5	32.9	1201	13 BX402419	BX402419 BX402419
14	1140	31.4	614	9 AW370693	AW370693 QV1-BT026
15	1137	31.3	613	9 AW410534	AW410534 fh06h06.x
16	1126	31.0	659	9 AW057736	AW057736 wx02h09.x
17	1115.5	30.7	1050	10 BF984645	BF984645 602309702
c 18	1097	30.2	670	9 AI906012	AI906012 RC-BT105-
19	1094	30.2	733	12 B1115219	B1115219 602861672
20	1073	29.6	588	2 HSM068717	Bx478931 Homo sapi
21	1063	29.3	563	10 BE297998	BE297998 601118462
22	1049	28.9	583	2 HSM068900	Bx479114 Homo sapi
23	1046	28.8	569	9 AL701765	AL701765 DKF2p686P
c 24	933	25.7	669	10 BF108852	BF108852 7164h02.x
25	915	25.2	641	10 BB621058	BB621058 BB621058
c 26	914	25.2	872	13 BX391706	BX391706 BX391706
27	893	24.6	474	9 AL701569	AL701569 DKF2p686A
28	880	24.3	561	12 BM720098	BM720098 UI-E-E00-
29	869	24.0	876	13 BQ768889	BQ768889 UI-M-F10-
c 30	847	23.3	514	10 BF998814	BF998814 QV0-GN014
31	842.5	23.2	657	10 BB620332	BB620332 BB620332
32	839.5	23.1	632	10 BB655484	BB655484 BB655484
33	830	22.9	539	14 CA544382	CA544382 C0648D06-
34	811	22.4	923	13 BQ88953	BQ88953 AGENCOURT
c 35	788.5	21.7	499	10 BF746234	BF746234 RCL-BT025
c 36	788.5	21.7	499	10 BF746304	BF746304 RCL-BT025
c 37	761	21.0	1057	12 BM907440	BM907440 AGENCOURT
c 38	747.5	20.6	486	9 AI906364	AI906364 RC-BT108-
39	743	20.5	447	10 BF436479	BF436479 7p13c10.x
c 40	738	20.3	995	13 BX336779	BX336779 BX336779
c 41	734	20.2	1201	13 BX402418	BX402418 BX402418
42	729	20.1	450	9 AI906050	AI906050 RC-BT105-
43	729	20.1	508	10 BE065758	BE065758 RC2-BT031
44	723.5	19.9	660	12 B1557977	B1557977 603236665
c 45	709	19.5	455	12 BG991986	BG991986 MR2-HT116

ALIGNMENTS

RESULT 1  
AK031099  
LOCUS  
DEFINITION  
AK031099 Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length enriched library, clone:5930404N10 product:v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian), full insert sequence.  
ACCESSION  
AK031099  
VERSION  
AK031099.1 GI:26082143  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
4323 bp mRNA linear HTC 05-DEC-2002

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
92279253  
PUBMED  
10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
PUBMED  
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
PUBMED  
11076861

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Rost, C., Schumacher, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
PUBMED  
11217851

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 4323)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/.

FEATURES  
Location/Qualifiers  
Source  
1..4323  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:5930404N10"  
/db\_xref="taxon:10090"  
/clone="5930404N10"  
/tissue\_type="forelimb"  
/dev\_stage="13 days embryo"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
misc\_feature  
48..3818  
/note="putative  
v-erb-b2 erythroblastic leukemia viral oncogene homolog 2,  
neuro/glioblastoma derived oncogene homolog (avian)  
(MGI:95410, GB|U71126, evidence: BLASTN, 99%,  
match=449)"  
BASE COUNT 933 a 1247 c 1226 g 917 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.03e-242 Length: 4323  
Score: 3087.50 Matches: 556  
Percent Similarity: 90.20% Conservative: 33  
Best Local Similarity: 85.15% Mismatches: 63  
Query Match: 85.10% Indels: 1  
Gaps: 1  
SEQ3 (1-653) x AK031099 (1-4323)  
Qy 1 MetGluLeuAlaAlaLeuCysArgTrrpGlyLeuLeuAlaLeuProGlyAla 20  
|||||  
Db 48 ATGGAGCTGGCGGCGCTGGTGGGGTTCCTCTCCGCTCTCCGCGGAGCC 107  
Qy 21 AlaserThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
|||  
Db 108 GCGGGTACCCCAAGTGTGTACCGGTACCGACATGAAGTTGCGACTCCCTGCCAGTCTGAG 167  
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
|||||  
Db 168 ACCACCTGGACATGCTTCCGACCTCTACGAGGCTGTGAGGTGGTGGCAATTTG 227  
Qy 61 GluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80  
|||||  
Db 228 GAGCTTACCTTACCTGGCGGCGCAATGCCAGGCTCTCATCTCTGAGGACATCCAGGAAGTC 287  
Qy 81 GlnGlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
|||||  
Db 288 CAGGGATACATGCTATCGCTCACACCGAGTGAACACCTCCACCTGCAGAGGTTGCGC 347  
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAspAsnGly 120  
|||||  
Db 348 ATCGTGAGAGGAGTCACTAGCTCTTTGAGGACAAAGTATGCCCTGGCTGTGTAGACACCGA 407  
Qy 121 AspProLeuAsnAsn---ThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlu 139  
|||||  
Db 408 GACCCCTTTGGACCAACGTCACCGCGCCCGCCGAGAGAACCGGCGCGGAGGCTCGCGGAG 467  
Qy 140 LeuGlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnPro 159

```

Db      |||||||TCCGAAGTCTCACAGATCTTGAAGGAGGAGTGTTCATCCGTGGGAACCTT 527
Qy      |||||||GlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeu 179
Db      |||||||TCTCTACAGAGATGTTTGTGGAGGATGCTCCCGTAAGAAATACACAGCTG 587
Qy      |||||||AlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCys 199
Db      |||||||GCTCCTGCGACATGCACACCAATCGTTCCCGGGCTGTCCACCTTGTGCCCAACCTGC 647
Qy      |||||||LysGlySerArgCysTrpGlySerSerGlnAspCysGlnSerLeuThrArgThrVal 219
Db      |||||||648 AAGACAATCACTGTGTGGGTGAGAGTCTCTGAAGACTGTCAAGTCTTGACCTGGCACCATC 707
Qy      |||||||220 CysAlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGln 239
Db      |||||||708 TGTACTAGTGTGCTGCCCGGTGCAAGGGCGGCTGCCCACCTGACTGTGTGCATGAGCAG 767
Qy      |||||||240 CysAlaLaglyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsn 259
Db      |||||||768 TGTGCTGCAGCTGCACGGGTCCCAAGCATTTCTGACTGACTGGCTGCTCCACTTCAAT 827
Qy      |||||||260 HisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPhe 279
Db      |||||||828 CATAGTGTATCTGTGAGCTGCACTGCCCGGCCCTCATCACTACACACAGACACCTTC 887
Qy      |||||||280 GluSerMetProAsnProGluCysLysGlyArgThrPheGlyAlaSerCysValThrAlaCys 299
Db      |||||||888 GAGTCCATGCTCAACCTCGAGGGTGTGTACACCTTTGGTGCAGCTGTGTGACCACCTGC 947
Qy      |||||||300 ProTyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsn 319
Db      |||||||948 CCTTACAACTACCTCTCCACGAAAGTGGATCCTGCACTCTGGTGTGCTGCCCGCAACAAC 1007
Qy      |||||||320 GlnGluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAla 339
Db      |||||||1008 CAAGAGGTTCACAGCTGAGGACGGAACACAGCGGTGTGAGAAATGAGCAAGCCCTGTGCT 1067
Qy      |||||||340 ArgValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAla 359
Db      |||||||1068 GGAGTATGCTATGCTGGGATGAGGACCTCCGAGGGGCGAGGCCATCACCAGTGAC 1127
Qy      |||||||360 AsnIleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGlu 379
Db      |||||||1128 AATATCCAGGAGTTGCTGGCTGCAAGAAGATCTTGGGAGCCTGGCATTTTTCGGGAG 1187
Qy      |||||||380 SerPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnVal 399
Db      |||||||1188 AGCTTTGATGGGAACCCCTCTCCGGCGTTCGCCCACTGTAAGCCAGAGCATCTCCAAGTG 1247
Qy      |||||||400 PheGluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTtpProAspSerLeu 419
Db      |||||||1248 TTGGNAACCTTGGAGGAGATCACAGTTTACTATACATTCAGATGATGCCAGAGAGCTTC 1307
Qy      |||||||420 ProAspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGly 439
Db      |||||||1308 CAAGACCTTCAGTCTTCCAGAACCTTCGGGTTCATTCGGGCGAGGATTCCTCATGATGTT 1367
Qy      |||||||440 AlaTyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArg 459
Db      |||||||1368 GCTTACTCATTTGAGCTGTGCAAGGCTGGGATTCACCTCACTCGGGGCTACGCTCACTGCGG 1427
Qy      |||||||460 GluLeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThr 479
Db      |||||||1428 GAGCTGGCACTGATGGCTCTCATTCACCGCAACCCATCTCTCTGTTGTAAACACT 1487
Qy      |||||||480 ValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArg 499
Db      |||||||1488 GTACCTTGGGACCACTCTTCCGGNACCCGACACCGCCCTACTCCAGTGGGAACCGG 1547
Qy      |||||||500 ProGluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHis 519

```

```

Db      1548 CCAGAAGAGGATGTGGTCTTGAAGGCTTGGTCTGTAACTCACTGTGTGCCGTGGGCAC 1607
Qy      520 CysTrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGlu 539
Db      1608 TGCTGGGGGCCAGGGCCACCAGTGTCTAACTGCAGTTCAGTTCTCCGGGGCCAGGAG 1667
Qy      540 CysValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHis 559
Db      1668 TGTGTGGAGGAGTCCGAGTATGGAAGGGCTCCCGAGGAGTATGTGAGGGGCAAGCAG 1727
Qy      560 CysLeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyPro 579
Db      1728 TGTCTGCCATGCCACCGAGTGTGAGGCTCAAAACAGCTCGGAGACCTGCTATGGAATCG 1787
Qy      580 GluAlaaspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArg 599
Db      1788 GAGGCTGACCACTGTGAGGCTTGTGCCACACACAGGACTCATCTTCCTGTGTGGCTCGC 1847
Qy      600 CysProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGlu 619
Db      1848 TGCCCCAGTGTGTGAAGCCAGACCTCTCTACATGCCCTATCTGGAAGTACCCGATGAG 1907
Qy      620 GluGlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAsp 639
Db      1908 GAGGCGATATGTCAGCCATGCCCATCACTCACTGACCCACTCATGTGTGGACCTGGACGAA 1967
Qy      640 LysGlyCysProAlaGluGlnArgAlaSerProLeuThr 652
Db      1968 CGAGGCTGCCAGAGAGAGAGAGCCAGCCAGCCAGTGACA 2006

RESULT 2
LOCUS      AK083669
DEFINITION Mus musculus 9 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:D030063B12 product:v-erb-b2 erythroblastic
leukemia viral oncogene homolog 2, neuro/glioblastoma derived
oncogene homolog (avian), full insert sequence.
ACCESSION  AK083669
VERSION     AK083669.1 GI:26101404
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus
ORGANISM    Mus musculus
REFERENCE   1
AUTHORS     Carninci, P. and Hayashizaki, Y.
TITLE       High-efficiency full-length cDNA cloning
JOURNAL     Meth. Enzymol. 303, 19-44 (1999)
MEDLINE     99279253
PUBMED      10349636
REFERENCE   2
AUTHORS     Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE       Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL     Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE     20499374
PUBMED      11042159
REFERENCE   3
AUTHORS     Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE       RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL     Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE     20530913
PUBMED      11076861
REFERENCE   4
AUTHORS     Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

```

Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schraml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Satai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Caldarelli, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustingich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
PUBMED  
11217851  
REFERENCE  
5

AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 4463)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, H., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitsu, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE	Journal of Experimental Animal Research
JOURNAL	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse-Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/ URL:http://fantom.gsc.riken.go.jp/

```

FEATURES
source
1. .4463
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:D030063B12"
/db_xref="taxon:10090"
/clone="D030063B12"
/tissue_type="whole body"
/dev_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="9 days embryo"
1. .4463
misc_feature
/note="v-erb-b2 erythroblastic leukemia viral oncogene
homolog 2, neuro/glioblastoma derived oncogene homolog
(avian) (MGDI|MG1:95410, GB|U71126, evidence: BLASTN, 99%,
match=449),"

```

[illegible]







	AUTHORS	Carninci,P. and Hayashizaki,Y.
	TITLE	High-efficiency full-length cDNA cloning
	JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
	MEDLINE	99279253
	PUBMED	10349636
	REFERENCE	2
	AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
	TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
	JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
	MEDLINE	20499374
	PUBMED	11042159
	REFERENCE	3
	AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Sumi,N., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,H., Itoch,M., Suno,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashivagi,K., Fujiwara,E., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M., Yoneda,Y., Isihkawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
	TITLE	Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer
	JOURNAL	Genom Res. 10 (11), 1757-1771 (2000)
	MEDLINE	20530913
	PUBMED	11076861
	REFERENCE	4
	AUTHORS	Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyouasa,H., Kondo,S., Yamanaka,I., Saio,T., Okazaki,Y., Gojohori,T., Bono.H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fieischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis.S., Matsuo,Y., Nikaido.I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki.R., Tomita,M., Wagner.L., Washio,T., Sakai,K., Okido,T., Furuno.M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga.N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult.C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill.D., Hofmann.M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons.P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring.B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki.H., Sato,K., Schonbach,C., Seya.T., Shibata.Y., Storch,K.F., Suzuki,H., Toyo-Oka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilming.L., Wynshaw-Boris.A., Yoshida.K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.
	TITLE	Functional annotation of a full-length mouse cDNA collection
	JOURNAL	Nature 409 (6821), 685-690 (2001)
	MEDLINE	21085660
	PUBMED	11217851
	REFERENCE	5
	AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
	TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
	JOURNAL	Nature 420, 563-573 (2002)
	REFERENCE	6 (bases 1 to 2662)
	AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno.M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiroaka,T., Horii,F., Imotani,K., Ishii.Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno.H., Kouda,M., Koya,S., Kuriharata,C., Matsuyama,T., Miyazaki,A., Nishi.K., Nomura,K., Numazaki.R., Onno.M., Okezaki,Y., Okido,T., Owa,C., Salto,H., Salto,R., Sakai.C., Sakai.K., Sano.H., Sasaki,D., Shibata,K., Shibata,Y., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki.H., Tagami,M., Tagawa,A., Takahashi.F., Tanaka,T., Teijima,Y., Toyta.T., Yamamura.T., Yasunishi,A., Yoshida.K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
	TITLE	Direct Submission
	JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC).
	REFERENCE	1



Db	1476	AAGGAANAATACAGCGCTTTTTCGTGATTCAGGCTGGCCTGTGATACTGACGTGACGTACCTCCAT	1535
Qy	424	ValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAlaValTrpSerLeu	443
Db	1536	GCTTTCGAGAACCTAGATAATAGTGTGGCAGACAAGACACATGGTCAGTTTCTTTG	1595
Qy	444	ThrLeuGlnGlyLeuGlyIleSerTrpLeuGlySerLeuArgGluLeuGlySer	463
Db	1596	CGGTCGTGGCTGAACATCACATCACTGGGCTGCGTTCCTCAAGGATCATGTCAT	1655
Qy	464	GlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrValProTrpasp	483
Db	1656	GGGATGTGATCATCTTCGGAAACCGAATTTGTGTACGCAACACAACTAAGCTGAAA	1715
Qy	484	GlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgProGluAspGlu	503
Db	1716	AAACTCTTCGGACACCCCAATCAGAAACCAACCAATCATGAACACAGAGCTGAGAAAC	1775
Qy	504	CysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaAargGlyHisCysTrpGlyPro	523
Db	1776	TGCAGGCGGTGAACACAGCTGTGCAATCTTTATGCTCGGAAGGCTGCTGGGCGCT	1835
Qy	524	GlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCysValGluGlu	543
Db	1836	GAGCCAGGAGTGTGTCTCTCCAGAAATGTGAGCAGAGGAGGAGTGGTGGAGAAA	1895
Qy	544	CysArgValLeuGlnGlyLeuProArgGluTrpValAsnAlaAargHisCysLeuProCys	563
Db	1896	TGCACATCCCTGGAGGGGGAACCAACAGGAGTTGTGGAAATTCGTGAATGATCCAGTGC	1955
Qy	564	HisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGluAlaAspGln	583
Db	1956	CATCCAGAATGCTGCCCGCCGCAATGAATCACCCTGTACAGCGAGGGACCCAGACAAC	2015
Qy	584	CysValAlaCysAlaHisTyrIlyAsnProProPheCysValAlaAargCysProSerGly	603
Db	2016	TGCATCCAGTGTGCCCATACATTATGTCGACACACATGTGTCAGACCTGCCAGCTGCG	2075
Qy	604	ValIlyProAspLeuSerTyrMetProIleTrpIlyPheProAspGluGluGlyAlaCys	623
Db	2076	ATCATGGGAGAGACAACACTGTG---GTCTGGAAAGTATGCAGATGCCAATAATGTCTGC	2132
Qy	624	GlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspIlyGlyCys	642
Db	2133	CACCTATGCCAGCCCACTGTACCTATGATGTGCTGGCCCGAGGCTCTCAAGGATGT	2189
RESULT 5			
AK004944			
LOCUS			
DEFINITION		2936 bp mRNA linear HTC 05-DEC-2002	
		Mus musculus adult male liver cDNA, RIKEN full-length enriched	
		library, clone:1300008123 product:epidermal growth factor receptor,	
		full insert sequence.	
ACCESSION	AK004944		
VERSION	AK004944.1	GI:12836511	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
JOURNAL	1		
MEDLINE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
PUBMED	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.		
	Normalization and subtraction of cap-trapper-selected cDNAs to		
	prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE			
AUTHORS	3		
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		
	Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,		
	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,		
	Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,		
	Fujimake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachihi, M.,		
	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,		
	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.		
	RIKEN integrated sequence analysis (RISA) system-384-format		
	sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE			
AUTHORS	4		
TITLE	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,		
	Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,		
	Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yananaka, R.,		
	Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R.,		
	Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,		
	Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,		
	Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,		
	Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,		
	Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,		
	Baldarelli, R., Barsh, G., Blake, J., Brownstein, M.J., Buit, C.,		
	Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Buit, C.,		
	Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,		
	Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,		
	Marchionni, L., Washima, J., Mazzarelli, J., Momberts, P., Nordone, P.,		
	Ring, B., Rindwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,		
	Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,		
	Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,		
	Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.		
	and Hayashizaki, Y.		
	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409 (6821), 685-690 (2001)		
MEDLINE	21085660		
PUBMED	11217851		
REFERENCE			
AUTHORS	5		
TITLE	The FANTOM Consortium and the RIKEN Genome Exploration Research		
	Group Phase I & II Team.		
	Analysis of the mouse transcriptome based on functional annotation		
	of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
REFERENCE			
AUTHORS	6 (bases 1 to 2936)		
TITLE	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,		
	Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,		
	Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,		
	Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,		
	Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,		
	Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,		
	Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,		
	Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,		
	Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,		
	Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,		
	Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,		
	Yoshino, M., Muramatsu, M. and Hayashizaki, Y.		
	Direct Submission		
TITLE	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of		
JOURNAL	Physical and Chemical Research (RIKEN), Laboratory for Genome		
	Exploration Research Group, RIKEN Genomic Sciences Center (GSC),		
	RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,		
	Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,		
	URL: http://genome-gsc.riken.go.jp/, Tel: 81-45-503-9222,		
	Fax: 81-45-503-9216)		
	Please visit our web site (http://genome.gsc.riken.go.jp/) for		
	further details.		
COMMENT			
	cDNA library was prepared and sequenced in Mouse Genome		
	Encyclopedia Project of Genome Exploration Research Group in Riken		
	Genomic Sciences Center and Genome Science Laboratory in RIKEN.		
	Division of Experimental Animal Research in Riken contributed to		
	prepare mouse tissues. First strand cDNA was primed with a primer		
	[5' GAGAGAGAGCGCGCACTGAGTGTGTGTGTGTGTGTGT 3'], cDNA was		
	prepared by using trehalose thermo-activated reverse transcriptase		











## COMMENT

Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

## FEATURES

source

Location/Qualifiers  
1..757

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLACE2000402"  
/tissue\_type="placenta"  
/clone\_lib="PLACE2"  
/note="Vector: pME18SFL3"

BASE COUNT 149 a 227 c 231 g 145 t 5 others

## ORIGIN

## Alignment Scores:

Pred. No.: 5,88e-102 Length: 757  
Score: 1364.00 Matches: 246  
Percent Similarity: 98.40% Conservative: 0  
Best Local Similarity: 98.40% Mismatches: 4  
Query Match: 97.60% Indels: 0  
Dbs: 9 Gaps: 0

SEQ3 (1-653) x AUI40362 (1-757)

QY 332 GluLysCysSerLysProCysAlaArgValCysTyrGlyLeuGlyMetGluHisLeuArg 351  
DB 3 GAGAAAGTCAGCAAGCCCTGCGCGAGTGCTATGCTGGGCATGGACACTTGGCA 62  
QY 352 GluValArgAlaValThrSerAlaAsnIleGlnGluPheAlaGlyCysLysIlePhe 371  
DB 63 GAGGTGAGGGCAGTTACCAAGTGCCCAATATCCAGGAGTTGTGGCTGCAAGAGATCTT 122  
QY 372 GlySerLeuAlaPheLeuProGluSerPheAspGlyAspProAlaSerAsnThrAlaPro 391  
DB 123 GGGAGCCCTGGCATTTCTCCGGAGAGCTTTCATGGGGAGCCAGCCCTCAACACTGCCCG 182  
QY 392 LeuGlnProGluGlnLeuGlnValPheGluThrLeuGluGluIleThrGlyTyrLeuTyr 411  
DB 183 CTCACCCAGAGCAGCTCCAAGTGTTCAGACTCTGGAGAGATCACAGGTACCTATAC 242  
QY 412 IleSerAlaIleProAspSerLeuProAspLeuSerValPheGlnAsnLeuGlnValIle 431  
DB 243 ATCTCAGCATGGCGGACAGCTGCTGACCTCAGCGTCTCCAGAACCTGCAAGTAATC 302  
QY 432 ArgGlyArgIleLeuHisAsnGlyAlaTyrSerLeuThrLeuGlnGlyLeuSer 451  
DB 303 CGGGAGCAATTCACAAATGGCCCTACTCGCTGACCTGCAAGGGCTGGGCATCAGC 362  
QY 452 TrpLeuGlyLeuArgSerLeuArgGluLeuGlySerGlyLeuAlaLeuIleHisAsn 471  
DB 363 TGGCTGGGGCTCGCTCACTGAGGAACTGGGCACTGGGCTCATCCACCATAAC 422  
QY 472 ThrHisLeuCysPheValHisThrValProTrpAspGlnLeuPheArgAsnProHisGln 491  
DB 423 ACCCACTCTGCTTCGTCACACAGCGTGGCCCTGGGACCACTTTCGGAACCCGACCAA 482  
QY 492 AlaLeuLeuHisThrAlaAsnArgProGluAspGluCysValGlyGluLeuAlaCys 511  
DB 483 GCTCTGCTCCACACTGCCAACCGCCAGAGGACGAGTGTGGGGAGGGCGCTGGCCCTGC 542  
QY 512 HisGlnLeuCysAlaArgGlyHisCysTrpGlyProGlyProThrGlnCysValAsnCys 531  
DB 543 CACCAGCTGTGGCCCGGGGACACTGCTGGGGTCCAGGGCCACCACCTGCTGCAACTGC 602

QY 532 SerGlnPheLeuArgGlyGlnGluCysValGluGluCysArgValLeuGlnGlyLeuPro 551  
DB 603 AGCCAGTTCTTCGGGGCCAGGANTGCGTGGAGGAATGCCGAGTACTGCANGGGCTCCC 662  
QY 552 ArgGluTyrValAsnAlaArgHisCysLeuProCysHisProGluCysGlnProGlnAsn 571  
DB 663 AGGAGTAGTATGAATGCCAGGCACCTGTTTGGCGTCCACCCCTGAGTGTGAGCCCGAGAT 722  
QY 572 GlySerValThrCysPheGlyProGluAla 581  
DB 723 GCCTCANTGACTGTGTTTGGACCGGANGCT 752

RESULT 9  
AK031542  
LOCUS  
DEFINITION  
Mus musculus 13 days embryo male testis cDNA, RIKEN full-length  
enriched library, clone:6030449F08 product:v-erb-b2 erythroblastic  
leukemia viral oncogene homolog 2, neuro/glioblastoma derived  
oncogene homolog (avian), full insert sequence.  
AK031542  
VERSION  
AK031542.1 GI:26327396  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
Carlini, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
PUBMED  
10349636  
2  
Carlini, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
PUBMED  
11042159  
3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Harada, A.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
PUBMED  
11076861  
4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakawa, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,  
Quackenbush, J., Schriml, L. M., Stauble, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bulc, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.  
and Hayashizaki, Y.

**TITLE**  
Functional annotation of a full-length mouse cDNA collection  
**JOURNAL**  
Nature 409 (6821), 685-690 (2001)  
**MEDLINE**  
21085660  
**PUBMED**  
11217851  
**REFERENCE**

**AUTHORS**  
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420 563-573 (2002)

**JOURNAL**  
6 (bases 1 to 3110)  
**REFERENCE**

**AUTHORS**  
Adachi,S., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,J., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Inotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Koima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,Y., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sugabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toyota,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

**TITLE**  
Direct Submission  
**JOURNAL**  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

**COMMENT**  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.go.jp/  
URL:http://fantom.gsc.riken.go.jp/.  
Location/Qualifiers  
1. .3110  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM.DB:6030449F08"  
/db\_xref="taxon:10090"  
/clone="6030449F08"  
/sex="male"  
/tissue.type="testis"  
/dev\_stage="Riken full-length enriched mouse cDNA library"  
/clone\_lib="Riken full-length enriched mouse cDNA library"  
/gene\_stage="13 days embryo"  
.L. 2648  
/note="unnamed protein product; putative v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian) (MGDI|GI:95410, GB|U71126, evidence: BLASTN, 99%, match=449)"  
/codon\_start=3  
/protein\_id="BAC27442.1"  
/db\_xref="GI:26327397"  
/translation="AFLPESFDGNSSGVAPLKPEHLQVFELEETGYLYISAWPES FDLVSDFQLRNPHIRGLHDGAYSLTLOGLGTHGLSRLEDLSGLAIRHNTLFCF VNTVPDOLFRNPAGLLHGNSRPPEACLEGVLNCISLCARGHCMPGTQCVCNCSOF LRGQCECRCKYWGKLPREYVRGKHCLPCHPGPCPONSETCYGSADQCEACAHYKD SSSVCARCPGWPKPDLSYPTWKYPDEEGICQPCINTCSCVDLDERGCPAPQRAS VTFIATVVGVLLFIIVVICIGILIRKEERQIKRYTMRLLETTELVEPLTPSGAVPN QAMRILKETLRKLKGDSGAFYVGKIWIIPDGENVAPIVAIKVIRENTSPKANKE ILDEAYVMAGVSPYSVRLGLCLSTVOLVTQMPYGCGLLDHVHRGRGLGSDLLN WCVOIAGMSYLEELVRHDLAAARNLVKPSNHVKINDPFGLARLLIDPETEHADGG KYPIEKMALESILRRFTHQDSWYSYGVTVWELMFGAKYPDGPAREIDPLEKGER LPQPCTIDTIVTWKCMWDSECRPELFREYSEFSRWDPQPRFVIONEDLGPPS PMDTFYSLBEDDDGEIYALVPOQGFPDPAITGSTAHHRRSHRSSARS GG GELTGFLSEEPSPRLAPSEGAEVDFDGLAVGVTGKGLSQSLPHDLSPLOYRSES

**SOURCE**  
Source

DPTLPPTTGDGVVAPLACSPQEVYVNQPEVRPQSPLTPGPPPPPIRPAGATLERPKT  
 LSPKNGVKVDVFAFGVGAVENPYLAIPRAGTASQPHPSPAFPAFNDLYWDCNSSDSBO  
 GPPPSFEETPTAENPEYLGLDVPV\*

BASE COUNT    677 a    886 c    891 g    656 t

ORIGIN

Alignment Scores:

Pred. No.:	5,35e-99	Length:	3110
Score:	1339.00	Matches:	235
Percent Similarity:	90.65%	Conservative:	17
Best Local Similarity:	84.53%	Mismatches:	26
Query Match:	36.91%	Indels:	0
DB:	11	Gaps:	0

SEQ3 (1-653) x AK031542 (1-3110)

```

QY      375 AlaPheLeuProGluSerPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnPro 394
       |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db      3 GCATTTTTCGGGAGAGCTTGATGTGAACCCCTCCTCGGCCTGTCGCCCACTGAAGCCA 62

QY     395 GluGlnLeuGlnValPheGluThrLeuGluGluIleThrGlyTyrlleSerAla 414
       |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db      63 GAGCATCTCCAAAGTGTCAAACCCCTGGAGGAGATCACAGGTTCACATATACATTTCAGCA 132

QY     415 TrpProAspSerLeuProAspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArg 434
       |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db     123 TGGCCAGAGAGCTTCCAAGACCCTCAGTGTCTTCCAGAACCCTTCGGGTGTCATTCGGGACGG 182

QY     435 IleLeuHisAsnGlyAlaTyrlerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGly 454
       |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db     183 ATTCTCCATGATGGTGCTTACTCATTTGACGTTTCAAGGCGCTGGGGATTCACTCACTGGGG 242

QY     455 LeuArgSerLeuArgGluLeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeu 474
       |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db     243 CTACGCTCACTCGGGAGGTGGAGCATGGATTGGCTCTCATTCACCGCAACACCCCATCTC 302

QY     475 CysPheValHisThrValProTpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeu 494
       |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db     303 TGC TTGTAAACACTGTACTTGGGACCAAGCATCTCCGGAACCCGACACGAGCCCTTACTC 362

QY     495 HisThrAlaAsnArgProGluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeu 514
       |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db     363 CACAGTGGGAACCGGCCAAGAGGCATGTGGTCTTGAGGGCTTGTCTGTAACCTCACTG 422

QY     515 CysAlaArgGlyHisCysTrpGlyProGlyProThrGlnCysValAsnCysSerGlnPhe 534
       |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db     423 TGTGCCGTGGGCACTGTCTGGGGCCAGGGCCCCACCCAGTGTGTCAACTGCAGTCAGTTTC 482

QY     535 LeuArgGlyGlnGluCysValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrl 554
       |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db     483 CTCGGGGCCAGAGTGTGTGGAGGAGTCCGAGTAGGAAGGGGCTCCCCAGGAGTAT 542

QY     555 ValAsnAlaArgHisCysLeuProCysHisProGluCysGlnProGlnAsnGlySerVal 574
       |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db     543 GTGAGGGGAACGACACTGTCTGCCATCCACCCCGAGTGTCAAGCTCAAAAACAGCTCGGAG 602

QY     575 ThrCysPheGlyProGluAlaaspGlnCysValAlaCysAlaHisTyrlrylsappProPro 594
       |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db     603 ACCTGCTATGGATCGGAGGCTGACCACTGTGAGGCTTGTGCCCACTACAAGGACTCATCT 662

QY     595 PheCysValAlaArgCysProSerGlyValLylysProAspLeuSerTyrlMetProIleTrp 614
       |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db     663 TCCTGTGTGGCTTCGTCGCCAGTGTGTGAAGCCAGACCTCTCCTACATGCCTATCTGG 722

QY     615 LysPheProAspGluGlyAlaCysGlnProCysProIleAsnCysThrHisSerCys 634
       |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db     723 AAGTACCCGGATGAGGAGGCATATGTGCACCATGCCCATCAACTGCACCCACTCATGT 782

QY     635 ValAspLeuAspAspIlyGlyCysProAlaGluGlnArgAlaSerProLeuthr 652
       |||.....|||.....|||.....|||.....|||.....|||.....|||.....
Db     783 GTGGACCTGGAGCAACGAGGCTGCCACGACAGACGAGAGCAGCCAGCTGATGACA 836

```

RESULT 10

```

CA489534
LOCUS       AGENCOURT_10810594 MAPcL Homo sapiens cDNA clone IMAGE:6722008 5',
            mRNA sequence.
ACCESSION   CA489534
VERSION     CA489534.1 GI:24952325
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Invitrogen Corp
            cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LRAM14284 row: a column: 16
            High quality sequence stop: 582.
FEATURES    Location/Qualifiers
             1..808
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:6722008"
             /cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1
             , LNCap"
             /lab_host="EMDH10B"
             /clone_lib="MAPcL"
             /note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
             Subtracted with brain, liver, lung, kidney and muscle.
             Directionally cloned. Priming method: oligo-dT. Average
             insert size: 1800 bp. Library amplification: 26,000 fold.
             Kristi A. Eglund, James J. Vincent, Robert Strausberg,
             Bungkok Lee & Ira Pastan: Discovery of new breast
             cancer genes encoding membrane and secreted proteins.
             Manuscript submitted."
BASE COUNT  169 a 255 c 225 g 159 t
ORIGIN
Alignment Scores:
Pred. No.:      2.28e-99      Length:      808
Score:          1333.00      Matches:    247
Percent Similarity: 94.80%      Conservative: 8
Best Local Similarity: 91.82%      Mismatches: 8
Query Match:    36.74%      Indels:     6
DB:             14          Gaps:       4

SEQ3 (1-653) x CA489534 (1-808)

QY 133 SerProGlyGlyLeuArgGluLeuGlnLeuArgSerLeuThrGluLeuLeuLysGlyGly 152
    |||||
Db 3 TCCCCAGGAGGCGCTCGGGAGCTGCAGCTTCGAGCGCTCACAGAGATCTTGAAGAGGAGG 62
    |||||
QY 153 ValLeuIleGlnArgAsnProGlnLeuGlnLeuGlnLeuGlnLeuGlnLeuGlnLeu 172
    |||||
Db 63 GTCCTGATCCAGCGGAACCCAGCTCTGCTACAGGACAGATTTTGTGAAGAGACATC 122
    |||||
QY 173 PheHisLysAsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCys 192
    |||||
Db 123 TTCCACAGAGAGGACCGCTGGCTCTACACTGATAGACACACCGCTCTCGGGCGCTGC 182
    |||||
QY 193 HisProCysSerProMetCysLysGlySerArgCysTrpGlyGluSerSerGluAspCys 212
    |||||
Db 183 CACCCTCTTCTCCGATGTGAAGGGCTCCCGCTGCTGGGGAGAGGGTCTCTGAGGATTGT 242

```

```

QY 213 GlnSerLeuThrArgThrValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuPro 232
    |||||
Db 243 CAGAGCCTGACGCGACACTGTCTGTCCCGTGGCTGTCCCGCTGCAAGGGGCCACTGGCC 302
    |||||
QY 233 ThrAspCysCysHisGlnGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCys 252
    |||||
Db 303 ACTGACTGCTGCCATGAGCAGTGTCTGCCGCTGCACGGGCCCCAAGCACTCTGACTGC 362
    |||||
QY 253 LeuAlaCysLeuHisPheAsnHisSerGlyIleCysGluLeuHisCysProAlaLeuVal 272
    |||||
Db 363 CTGGCTGCCCTCCCACTTCAACACAGTGGCATCTGTGAGCTGCATGCCAGCCCTGGTC 422
    |||||
QY 273 ThrTyrAsnThrAspThrPheGluSerMetProAsnProGluGlyArgTyrThrPheGly 292
    |||||
Db 423 ACCTACACACAGACACGCTTTGAGTCCATGCCAATCCAGAGGGCGGTATACATTCCGC 482
    |||||
QY 293 AlaSerCysValThrAlaCysProTyrAsnTyrLeuSerThrAspValGlySerCysThr 312
    |||||
Db 483 GCCAGCTGTGACTGCTGCTCCCTACAACTACCTTTCTAGCGAGCTGGGATCCTGCACC 542
    |||||
QY 313 LeuValCysProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgCysGlu 332
    |||||
Db 543 CTGCTCTCCCCCTGCACACAGAGGTGACAGCAGAGATGGGAACACACGGGTGTGAG 602
    |||||
QY 333 LysCysSerLysProCysAlaArgValCysTyrGlyLeuGlyMetGluHisLeuArgGlu 352
    |||||
Db 603 AAGTGCACAGCCCTGTGCCGAGTGTCTATGCTGGGCATGGGACACTGGGAGAG 662
    |||||
QY 353 ValArgAlaValThrSerAlaAsnIleGln-GluPheAla-GlyCysLysLys---IleP 371
    |||||
Db 663 GTGAGGGCAGTTACGAGTGCCCAATATCCAGAGAGTTTGCCTGGCTGCAAAAAGATCTTT 722
    |||||
QY 371 heGlySerLeuAlaPheLeuPro---GluSerPheAspGlyAspPro---AlaSerAsnT 389
    |||||
Db 723 GGGGAGCCTGCAGATTTCTGCCCGGAAACTTTTGTATGGGGAACCAACCTCCCAACA 782
    |||||
QY 389 hrAlaPro---LeuGlnProGlu 395
    |||||
Db 783 CTGGCCCCCGGTTTCCAGGCCCAAA 805
    |||||
RESULT 11
BE746725
LOCUS      BE746725
DEFINITION 601579159p1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3927927 5',
            mRNA sequence.
ACCESSION   BE746725
VERSION     BE746725.1 GI:10160717
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   NIH-MGC http://mgc.nci.nih.gov/.
            1 (bases 1 to 894)
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: DCTD/DTF
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM758 row: e column: 16
            High quality sequence start: 3
            High quality sequence stop: 762.
FEATURES    Location/Qualifiers
             1..894
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:3927927"

```

/tissue\_type="adenocarcinoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_9"

/note="Organ: ovary; Vector: pORF7; Site:1; XhoI; Site:2;  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGACGAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 172 a 285 c 252 g 185 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 5,11e-96 Length: 894  
 Score: 1293.00 Matches: 261  
 Percent Similarity: 85.76% Conservative: 4  
 Best Local Similarity: 84.47% Mismatches: 29  
 Query Match: 35.64% Indels: 17  
 DB: 10 Gaps: 6

SEQ3 (1-653) x BE746725 (1-894)

QY 230 ProLeuProThrAspCysCysHisGluGlnCysAlaAlaGlyCysThrGlyProLysHis 249  
 DB 2 CCACTGCCACTGACTGCTGCCATGACAGCTGTGCGCGTGCACGGGCCCAAGCAC 61  
 QY 250 SerAspCysLeuAlaCysLeuHisPheAsnHisSerGlyLeuGlnCysPro 269  
 DB 62 TCGACTGCCCTGCCCTCCACTTCAACCACAGTGGCATCTGTGAGTGCACCTGCCA 121  
 QY 270 AlaLeuValThrTrpAsnThrAspThrPheGluSerMetProAsnProGluGlyArgTyr 289  
 DB 122 GCCCTGGTCACTACACACAGACAGCTTTGAGTCCATGCCCAATCCCGAGGCCGGTAT 181  
 QY 290 ThrPheGlyAlaSerCysValThrAlaCysProTyrAsnTyrLeuSerThrAspValGly 309  
 DB 182 ACATTGGCGCCAGCTGTGTGACTGTGCTGCCCTACAACTACCTTTCTACGAGCTGGGA 241  
 QY 310 SerCysThrLeuValCysProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGln 329  
 DB 242 TCTGCACTGCTGTGCCCTTGCACACACAGAGGTGACAGCAGGATGGAACACAG 301  
 QY 330 ArgCysGluLysCysSerLysProCysAlaArgValCysTyrGlyLeuGlyMetGluHis 349  
 DB 302 CGGTGTGAGAAGTGCAGACAGCCCTGTGCCGAGTGTGTATGCTGGCATGGAGCAC 361  
 QY 350 LeuArgGluValArgAlaValThrSerAlaAsnIleGlnGluPheAlaGlyCysLysLys 369  
 DB 362 TTGCGAGAGTGGGGCAGTTACAGTGCCTGCAATATCCAGAGTTTGTGCTGCTCAAGAAG 421  
 QY 370 IlePheGlySerLeuAlaPheLeuProGluSerPheAspGlyAspProAlaSerAsnThr 389  
 DB 422 ATCTTTGGGAGCTGGCATTTCTGCCGAGAGCTTTGATGGGACCCAGCTCCACACT 481  
 QY 390 AlaProLeuGlnProGluGlnLeuGlnValPheGluThrLeuGluGluLeuThrGlyTyr 409  
 DB 482 GCCCGCTGCCAGCAGCAGCTCCAAAGTGTGTTGAGACTCTGGAAGATCACAGTTAC 541  
 QY 410 LeuTyrIleSerAlaTrpProAspSerLeuProAspLeuSerValPheGlnAsnLeuGln 429  
 DB 542 CTATACATCTCAGCATGGCGGACAGCTGCTGCCCTGAGCTGACCTTCCAGAACCTGCMA 601  
 QY 430 ValIleArgGlyArgIleLeuHisAsnGlyAlaTyrSerLeuThrLeuGlnGlyLeuGlu 449  
 DB 602 GTATCCGGGGAGCAATCTGCACATGGCGCTACTCGCTGACCTGCAAGTGTGGG 661  
 QY 449 YlleSer-TrpLeuGlyLeuArgSerLeuArg---GluLeu-GlySerGlyLeuAlaLeu 467  
 DB 662 CATCAGTTCGGTGGGCTCGCTCACTTGCAGTGGACACTTGGGGCAGTGGACGTGGCC--- 718  
 QY 468 IleHisAsnThrHisLeuCysPheValHisThrValProTrpAspGlnLeuPheArg 487

Db 719 TCATCATAATAAACCCACTC-TGTTGTCGCCACCGTGCCTGGACCACTCTTCGG- 776  
 QY 488 AsnProHisGlnAlaLeuLeuHisThrAlaAsnArgProGluAspGluCysValGlyGlu 507  
 Db 777 AACCCAGCAGCTCT---GTCCACATTGCCAACCCGCAGAGGACC----- 818  
 QY 508 GlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCysTyr-----GlyProGlyPro 525  
 Db 819 ---ATTTTTGCAGGCCCTTGTGTCGCCACAGTGTCCCGTGTGTTGGTCCGCCCC 875  
 QY 526 ThrGlnCysValAsnCysSer 532  
 Db 876 ---CCGTGTCGTGGCGGCTCT 893  
 RESULT 12  
 LOCUS CA328613 795 bp mRNA linear EST 27-NOV-2002  
 DEFINITION UI-M-FY0-cda-e-16-0-UI.r1 NIH\_BMAP\_FY0 Mus musculus cDNA clone  
 IMAGE: 6826841 5', mRNA sequence.  
 ACCESSION CA328613  
 VERSION CA328613.1 GI:24546711  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 795)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaaps-remail.nih.gov  
 Tissue Procurement: Dr. Jim Lin, University of Iowa  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)  
 Seq primer: pYX-5.  
 Location/Qualifiers  
 1..795  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE: 6826841"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="NIH\_BMAP\_FY0"  
 /note="Organ: Brain; Vector: pYX-Asc; Site:1: EcoR I;  
 Site:2: Not I; The library was constructed according  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated  
 with EcoR I adaptor, digested with NotI and then cloned  
 directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is AGCGACACAG. This library was created for the University  
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
 Developing Mouse Nervous System', supported by National  
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
 program coordinator."  
 BASE COUNT 174 a 238 c 216 g 165 t 2 others  
 ORIGIN  
 Alignment Scores:

```

Pred. No.: 6.63e-89 Length: 795
Score: 1205.50 Matches: 219
Percent Similarity: 87.55% Conservative: 13
Best Local Similarity: 82.64% Mismatches: 31
Query Match: 33.23% Indels: 2
DB: 14 Gaps: 1

SEQ3 (1-653) x CA328613 (1-795)

QY 86 IleAlaHisAsn-GlnValArgGlnValProLeuGlnArgLeuArgGlyLeuValArgGlyTh 105
Db 3 ATCGGTCAACACGAGTGAACACAGTCCACATGCGAGAGTGGCATCGTGAGAGGAC 62
QY 105 rGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAspProLeuAsnAs 125
Db 63 TCAGCTCTTGGAGACAGATATGCTCCCTGGCTGTGTAGACACACGAGACCCCTTGACAA 122
QY 125 N---ThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGlnLeuArgSe 144
Db 123 CGTCACACCGCGCCGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 182
QY 144 rLeuThrGluLeuLeuLysGlyValLeuLeuLeuGlnArgAsnProGlnLeuLysGlyG 164
Db 183 TCTCACAGAGATCTTGAAGGGAGGAGTGTGATCGCTGGGAACCCCTCAGCTCTGCTACCA 242
QY 164 nAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThrLeuI 184
Db 243 GGACATGTTTGTGGAGGAGTGTCTCGTGAAGATTAACCACTGGCTCTGCTGCGACAT 302
QY 184 eAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGlySerArgCy 204
Db 303 GGACACCAATCTTCCGGGCTGTCCACCTTGTGCCCAACCTGCAAGACAAATCACTG 362
QY 204 STpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAlaGlyCy 224
Db 363 TTGGGGTGAGAGTCTTGAAGACTGTAGACTTTGACTGGCACCATCTGTACTAGTGGCTG 422
QY 224 sAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAlaAlaGlyCy 244
Db 423 TCCCCGGTGCAAGGCGCGCTGCCACCTGCTGTGCCATGAGCAGTGTGTGCGAGGCTG 482
QY 244 sThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGlyIleCy 264
Db 483 CACGGGTCCCAAGCATCTTGACTGTGCTGCTGCCCTGCCATTCATCATATAGTGTATCTG 542
QY 264 sGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSerMetProAs 284
Db 543 TGAGTGTCACTGCCCGGCGCTCATCACCCTACACACACAGACACCTTCGAGTCCATGCTCAA 602
QY 284 nProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsnTyrLe 304
Db 603 CCCTGAGGTCGCTACACTTTGGTGGCAGCTGTGTGACCACTGCCCTGCCCTACAACTACT 662
QY 304 uSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGlnGluValThrAl 324
Db 663 CTCACGGAAGTGGGATCTGCACCTCTGTCTGTCTGCCCGCAACACCAAGAGGTCAACAG 722
QY 324 aGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArgValCysTyrG 344
Db 723 TGAGACCGGAACACAGCGGTGTGAGAAATGCAAGCAAGCCCTGTGCTGGAGTATGCTATGG 782
QY 344 yLeuGlyMetGlu 348
Db 783 TCTGGGCATGGAG 795

RESULT 13
BX402419
LOCUS BX402419 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1032YB05 5-PRIME, mRNA sequence.
ACCESSION BX402419
VERSION BX402419.1 GI:30632074
KEYWORDS EST.

```

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8568.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSIAI0082H03QPlccluster=8568.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSIAI0082H03QPl.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1032YB05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 235 a 376 c 321 g 223 t 46 others
ORIGIN
Alignment Scores:
Pred. No.: 9.75e-88 Length: 1201
Score: 1194.50 Matches: 263
Percent Similarity: 52.18% Conservative: 0
Best Local Similarity: 52.18% Mismatches: 11
Query Match: 32.92% Indels: 231
DB: 13 Gaps: 3
SEQ3 (1-653) x BX402419 (1-1201)
QY 1 MetGluLeuAlaAlaLeuCysArgTrp-GlyLeuLeuAlaLeuLeuProGlyAl 20
Db 192 ATGGAGCTGGCGGCTTGTGCGCTGGYKGTCTCTCGCTTGTGCGCGGAGC 251
QY 20 aAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGl 40
Db 252 CGCGAGCACCAAGTGTGCACCGCACAGATGAAGCTGGCGGCTCCCTGCCAGTCCCGA 311
QY 40 uThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLe 60
Db 312 GACCCACCTGGACATGCTCCGCCACCTCTACCAGGCTGCCAGGTGGTGGAGGAACCT 371
QY 60 uGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVa 80
Db 372 GGAATCACTACCTACCTGCCCAATGCCAGCTGTCTCTCTGCA-GAC----- 419
QY 80 lGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuAr 100
Db 419 ----- 419
QY 100 gIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGl 120
Db 419 ----- 419
QY 120 yAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLe 140
Db 419 ----- 419

```

```

Qy 140 uGlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProG1 160
Db 419 -----
Qy 160 nLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAl 180
Db 419 -----
Qy 180 aLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLy 200
Db 419 -----
Qy 200 sGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCy 220
Db 419 -----
Qy 220 sAlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCy 240
Db 419 -----
Qy 240 sAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHi 260
Db 419 -----
Qy 260 sSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheG1 280
Db 419 -----
Qy 280 uSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPr 300
Db 419 -----
Qy 300 oTyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnG1 320
Db 420 ----AACTACTCTTCTACGACGCTGGGATCTCGACCCCTCGTGGCCCTGCACACCA 475
Qy 320 nGluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaAr 340
Db 476 AGAGTGACACGACGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGCAGCCCTGTGCCG 535
Qy 340 qValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAs 360
Db 536 AGTGTCTATGGCTGGCATGAGACACTTGCAGAGAGGTGAGGACAGTACCAGTGCCAA 595
Qy 360 nIleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSe 380
Db 596 TATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGGAGCTTGGCATTTCTGCCGGAGAG 655
Qy 380 rPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPh 400
Db 656 CTTTGTATGGGACCCAGCCTCCAACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT 715
Qy 400 eGluThrLeuGluGluThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPr 420
Db 716 TGAGACTCTGGAGAGATCACAGTTACCTATACATCTCAGCATGCGCGACACGCTGCC 775
Qy 420 oAspLeuSerValPheGlnAsnLeuGlnValIleArg-GlyArgIleLeuHisAsnGlyA 440
Db 776 TGACCTCAGGCTCTCCAGAACCTGCAAGTAATCCGGGGGACGAATTCGCACAAATGGCG 835
Qy 440 laTyrSerLeuThrLeuGln-GlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArg 459
Db 836 CCTACTCGCTGACCTGCAAGGCTGGGGCATCAGCTGGCTGGGGCTCGCTCCTCAGTGG 895
Qy 460 GluLeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThr 479
Db 896 GAATGGGCTAGTGACTGGCCCTCATCCACCATTAACACCACCTCTGCTTCTGGGCACAG 955
Qy 480 ValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArg 499
Db 956 GTGCCCTGGGACARCTCTTTC---GGNACCGCACAGCTCTGTC---ACATGCCAACGG 1009
Qy 500 ProGlu 501

```

```

Db 1010 CCAGAG 1015
RESULT 14
AW370693
LOCUS
DEFINITION
QV1-BT0260-011199-024-all BT0260 Homo sapiens cDNA, mRNA sequence.
ACCESSION
AW370693
VERSION
AW370693.1 GI:6875347
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 614)
HGCP http://www.ludwig.org.br/ORESTES.
AUTHORS
The FAPESP/LICR Human Cancer Genome Project
TITLE
The FAPESP/LICR Human Cancer Genome Project
JOURNAL
Unpublished
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV1&t2=QV1-BT0260-
011199-024-all&t3=1999-11-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 12
High quality sequence stop: 612.
FEATURES
source
Location/Qualifiers
1..614
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BT0260"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 119 a 198 c 172 g 124 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1,09e-83 Length: 614
Score: 1140.00 Matches: 197
Percent Similarity: 98.99% Conservative: 0
Best Local Similarity: 98.99% Mismatches: 2
Query Match: 31.42% Indels: 0
DB: 9 Gaps: 0
SEQ3 (1-653) x AW370693 (1-614)
Qy 153 ValLeuIleGlnArgAsnProGlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIle 172
Db 17 GTCTTGATCCAGCGGAACCCCGACGCTGCTACCAGACACGATTTGTGGAAGGACATC 76
Qy 173 PheHisLysAsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCys 192
Db 77 TTCACAAGAACAACACCGCTGGCTCTACACTGATAGACACCAACCGCTCTCGGGCCTGC 136
Qy 193 HisProCysSerProMetCysLysGlySerArgCysTrpGlyGluSerSerGluAspCys 212
Db 137 CACCCCTGTTCTCCGATGTGAAGGCTCCCGCTCTGGGGAGAGAGTTCTGAGGATGT 196
Qy 213 GlnSerLeuThrArgThrValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuPro 232

```

```

197  CAGAGCCTGACGGCACTCTCTGCGGTGGTGTGCGCGTGAAGGGCGCACTGCGCC 256
233  ThrAspCysCysHisGluGlnCysAlaAlaGlyCysThrGlyProIlyHisSerAspCys 252
257  ACTGACTGCTGCGCATGAGCAGTGTGTCGGGTGTCACGGGCCCAAGCAGCTGACTGC 316
253  LeuAlaCysLeuHisPheAsnHisSerGlyIleCysGluLeuHisCysProAlaLeuVal 272
317  CTGGCTGCTCCACTTCAACACACAGTGGCATCTGTGAGCTGCAGTGCACCGCCCTGGTC 376
273  ThrTyrAsnThrAspThrPheGluSerMetProAsnProGluGlyArgTyrThrPheGly 292
377  ACCTACACACAGACACGTTTGTAGTCCATGCCCAATCCCGAGGGCGGTATACATTCGGC 436
293  AlaSerCysValThrAlaCysProTyrAsnTyrLeuSerThrAspValGlySerCysThr 312
437  GCAGCTGTGTGACTGCTCCCTTACAACTACCTTCTACGAGCGTGGATCCCTGCACC 496
313  LeuValCysProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgCysGlu 332
497  CTCGTCTGCCCCCTGCACAAAGAGGTGACAGCAGAGGATGGAACACAGCGGTGGG 556
333  LysCysSerLysProCysAlaArgValCysTyrGlyLeuGlyMetGluHisLeuArg 351
557  AAGTGCAGCAAGCCCTGTTGCCGAGTGTCTATGTTCTGGGATGGAGCACTTGGGA 613

```

```

RESULT 15
AW410534
LOCUS
DEFINITION
  AW410534
  mRNA sequence.
  613 bp mRNA linear EST 29-JUN-2000
  fh06h06.x1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2961635 5',
  mRNA sequence.
ACCESSION
  AW410534
VERSION
  AW410534.1 GI:6936075
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 613)
  NIH-MGC http://mgi.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgabbs-remail.nih.gov
  Tissue procurement: ATCC
  CDNA Library Preparation: Ling Hong/Rubin Laboratory
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: National Institutes of Health Intramural
  Sequencing Center (NISC)
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Plate: LLCN56 row: O column: 12
  Seq primer: -21MI3 forward primer (ABI).
  Location/Qualifiers
    1..613
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:2961635"
      /tissue_type="rhodomyosarcoma"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH_MGC_17"
      /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
      Site_2: XhoI; cDNA made by oligo-dr priming.
      Directionally cloned into EcoRI/XhoI sites using the
      following 5' adaptor: GGCAGGAG(G). Size-selected >500bp
      for average insert size 1.8kb. Library constructed by
      Ling Hong in the laboratory of Gerald M. Rubin (University
      of California, Berkeley) using ZAP-cDNA synthesis kit
      (Stratagene) and Superscript II RT (Life Technologies)."
```

FEATURES

source

BASE COUNT

109 a 205 c 185 g 113 t 1 others

```

ORIGIN
Alignment Scores:
Pred. No.: 1,91e-83 Length: 613
Score: 1137.00 Matches: 195
Percent Similarity: 97.99% Conservative: 0
Best Local Similarity: 97.99% Mismatches: 4
Query Match: 31.34% Indels: 0
DB: 9 Gaps: 0
SQ03 (1-653) x AW410534 (1-613)
QY 455 LeuArgSerLeuArgGluLeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeu 474
Db 5 CTGCGCTCACTGAGGGAAGTGGGAGTGGAGTGGCCCTCATCCACCAACACACACCTC 64
QY 475 CysPheValHisThrValProTyrAspGlnLeuPheArgAsnProHisGlnAlaLeuLeu 494
Db 65 TGCTTCGTGCACACACGGGCCCTGGGACAGCTCTTTTCGGAACCCCGACCAAGCTCTGCTC 124
QY 495 HisThrAlaAsnArgProGluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeu 514
Db 125 CACACTGCCAACCGGCCAGAGAGTGTGTGGCGAGGGCTGGCTGCCACCAAGCTG 184
QY 515 CysAlaArgGlyHisCysTyrProGlyProGlyProThrGlnCysValAsnCysSerGlnPhe 534
Db 185 TCGCCCCGAGGGCACTGCTGGGGTCCAGGGCCACCCAGTGTGTCAACTGCAGCCAGTTC 244
QY 535 LeuArgGlyGlnGluCysValGluCysArgValLeuGlnGlyLeuProArgGluTyr 554
Db 245 CTTCGGGCCAGGAGTGGGTGGAGGAATGCCAGTACTGCAGGGGCTGCCACAGGAGTAT 304
QY 555 ValAsnAlaArgHisCysLeuProCysHisProGluCysGlnProGlnAsnGlySerVal 574
Db 305 GTGAATGCCAGCACTGTTTGGCGTGGCCACCTGAGTGTACGCCCCCAAGTGGCTCAGTG 364
QY 575 ThrCysPheGlyProGluAlaAspGlnCysValAlaCysAlaHisTyrLysAspProPro 594
Db 365 ACCCTGTTTGGACCGGAGGCTGACCAGTGTGTGGCTGTGGCCACTATAAGGACCCCTCCC 424
QY 595 PheCysValAlaArgCysProSerGlyValLysProAspLeuSerTyrMetProIleTrp 614
Db 425 TTCTGCGTGGCCCGCTGCCCCAGCGGTGTGAAACCTGNACCTCTCTACATGCCCATCTGG 484
QY 615 LysPheProAspGluGluGlyValaCysGlnProCysProIleAsnCysThrHisSerCys 634
Db 485 AAGTTTCCAGATGAGGAGGGCGCATGCCAGCTTGGCCCCCATCACTGCACCCACCTCTGT 544
QY 635 ValAspLeuAspAspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 545 GTGGACCTGGATGACAAGGGGTGCCCGCCGAGCAGAGAGAGCCGCTCTGACGCTCC 601

```

Search completed: October 16, 2003, 17:04:35  
Job time : 5138.68 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 16, 2003, 11:08:07 ; Search time 675.009 Seconds  
(without alignments)  
2540.503 Million cell updates/sec

Title: SEQ3

Perfect score: 3628

Sequence: 1 MELAALCRWGLLLALLPPGA.....CVDLDRKCPAERASPLTS 653

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEI=frame+p2n\_model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/HOLLERAN480/runat\_15102003\_131915\_20562/app\_query.fasta\_1.4685  
-DB=PublishedApplications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-NAXLEN=2000000000 -USER=HOLLERAN480 -CGN\_1\_1\_397 @runat\_15102003\_131915\_20562  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA.\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3628	100.0	3765	12	US-10-207-498-5 Sequence 5, Appli

2	3628	100.0	3768	9	US-09-811-123-8	Sequence 8, Appli
3	3628	100.0	3768	9	US-09-811-115-2	Sequence 2, Appli
4	3628	100.0	3768	10	US-09-854-356-9	Sequence 9, Appli
5	3628	100.0	3768	10	US-09-930-125-1	Sequence 1, Appli
6	3628	100.0	3768	12	US-10-313-644-1	Sequence 1, Appli
7	3628	100.0	4473	11	US-09-441-411-5	Sequence 5, Appli
8	3628	100.0	4473	12	US-10-101-510-81	Sequence 81, Appli
9	3628	100.0	4473	14	US-10-146-473-32	Sequence 32, Appli
10	3628	100.0	4473	14	US-10-207-655-44	Sequence 44, Appli
11	3628	100.0	4530	10	US-09-877-177-11	Sequence 11, Appli
12	3628	100.0	4530	12	US-10-007-926A-119	Sequence 119, App
13	3628	100.0	4530	12	US-10-101-510-124	Sequence 124, App
14	3628	100.0	4530	12	US-10-338-730-1	Sequence 1, Appli
15	3628	100.0	4530	14	US-10-177-293-125	Sequence 125, App
16	3628	100.0	4606	12	US-09-971-392-70	Sequence 70, Appli
17	3628	100.0	4642	14	US-10-198-846-10896	Sequence 10896, A
18	3628	100.0	9274	9	US-09-811-123-7	Sequence 7, Appli
19	3628	100.0	9274	9	US-09-811-115-1	Sequence 1, Appli
20	3585	98.8	4543	10	US-09-769-508-1	Sequence 1, Appli
21	3106	85.6	3955	10	US-09-870-759-117	Sequence 117, App
22	3106	85.6	3955	10	US-09-854-356-10	Sequence 10, Appli
23	3106	85.6	3955	12	US-09-751-708A-117	Sequence 117, App
24	3100.5	85.5	3771	10	US-09-854-356-11	Sequence 11, Appli
25	1644.5	45.3	2091	9	US-09-821-883-9	Sequence 9, Appli
26	1630.5	44.9	2070	9	US-09-821-883-7	Sequence 7, Appli
27	1614.5	44.5	1692	9	US-09-821-883-8	Sequence 8, Appli
28	1612	44.4	1665	9	US-09-821-883-6	Sequence 6, Appli
29	1608	44.3	1437	9	US-09-821-883-10	Sequence 10, Appli
30	1587	43.7	867	9	US-09-821-883-24	Sequence 24, Appli
31	1532.5	42.2	3633	9	US-09-725-433-1	Sequence 1, Appli
32	1532.5	42.2	5532	12	US-10-007-926A-137	Sequence 137, App
33	1532.5	42.2	5532	12	US-10-101-510-95	Sequence 95, Appli
34	1530.5	42.2	2643	14	US-10-172-620-15	Sequence 15, Appli
35	1530.5	42.2	10058	10	US-09-974-298-98	Sequence 98, Appli
36	1518.5	41.9	5264	10	US-09-920-300A-1731	Sequence 1731, Ap
37	1518.5	41.9	5264	12	US-10-099-926-1731	Sequence 1731, Ap
38	1518.5	41.9	5264	13	US-10-033-528-1731	Sequence 1731, Ap
39	1451	40.0	5484	10	US-09-940-101-1	Sequence 1, Appli
40	1451	40.0	5484	14	US-10-207-655-46	Sequence 46, Appli
41	1447	39.9	2601	10	US-09-940-101-3	Sequence 3, Appli
42	1435	39.6	4026	12	US-10-207-498-1	Sequence 1, Appli
43	1435	39.6	4879	12	US-10-101-510-137	Sequence 137, App
44	1435	39.6	4879	12	US-10-388-410-6	Sequence 6, Appli
45	1435	39.6	4879	14	US-10-172-620-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1

US-10-207-498-5  
; Sequence 5, Application US/10207498  
; Publication No. US20030143568A1  
; GENERAL INFORMATION:  
; APPLICANT: Elizabeth Singer  
; APPLICANT: Ralf Landgraf  
; APPLICANT: Dennis J. Slamon  
; APPLICANT: David Eisenberg  
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING  
; TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HEREULIN AND HER3  
; FILE REFERENCE: 30448.103-US-01  
; CURRENT APPLICATION NUMBER: US/10/207,498  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 60/308,431  
; PRIOR FILING DATE: 2001-07-27  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 3765  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(3765)

US-10-207-498-5

## Alignment Scores:

Pred. No.: 0 Length: 3765  
Score: 3628.00 Matches: 653  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

SEQ3 (1-653) x US-10-207-498-5 (1-3765)

Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGluYala 20  
Db 1 ATGAGAGTGGCGGCTTGTGCGGCTGGGGGCTCCTCGCCCTTGGCCCGGAGCC 60  
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
Db 61 GCGAGCACCAAGTGTGACCGGCACAGATGAAGCTGGGGCTCCTGCGAGTCCCGAG 120  
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
Db 121 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGTGGTGCAGGGAACCTG 180  
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
Db 181 GAACCTACCTACCTGCGCCACCAATGCCAGGCTGTCTTCTGCAAGGATATCCAGGAGGTG 240  
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
Db 241 CAGGGCTACGTGCTTCATCGCTCAACCAAGTGAAGGAGGCTCCACCTGCGAGGCTCGG 300  
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
Db 301 ATTGTGCGAGCACCCAGCTCTTTGAGACAACATATGCCCTGGCGCTGTAGACAATGGA 360  
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
Db 361 GACCCGTGAAATACACCCCTGTCTACAGGGGCTCCCGCAGGAGGCTGCGGGAGCTG 420  
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160  
Db 421 CAGCTTGAAGCTTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGGGAACCCCGAG 480  
Qy 161 LeuCystyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180  
Db 481 CTCTGCTACCCAGGACAGATTTGTGAAGGACATCTTCCACAAGAAACAACACAGCTGGCT 540  
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGGCTTCCACCCCTGTCTCCGATGTGTAAG 600  
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
Db 601 GGCTCCCGCTGCTGGGAGAGAGTCTTGAGGATTTGTGAGAGCTGACGCGCAGCTGTCTGT 660  
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
Db 661 GCCGGTGGCTGTGCCCTGCAAGGGGCCACTGCCACTGACTGCTGCCATGAGCAGTGT 720  
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
Db 721 GCTGCCGGCTGCAGGGGCCCAAGCACTCTGACTGCTGGCTGCCCTGCCCTTCAACAC 780  
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
Db 781 AGTGGCATGTGAGCTGCACTGCCAGCCCTGTGTCACCTACACACAGACAGCTTTGAG 840  
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
Db 841 TCCATGCCCAATCCGAGGGCCGGTATACATTCGGGCCAGCTGTGTGACTGCTGTGCC 900  
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320

Db 901 TACAACCTACCTTTCTACGGAGCTGGGATCTGCAACCCTCGTCTGCCCTCCACAACCAA 960  
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
Db 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1020  
Qy 341 ValCystyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
Db 1021 GTGTGCTATGGTCTGGGCATGGAGCACTTGGAGAGGTGAGGAGCTTACCAGTGCCCAAT 1080  
Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
Db 1081 ATCCAGGAGTTGTGTGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC 1140  
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
Db 1141 TTTGATGGGACCCAGCCTCCACACTGCCCGCTCCAGGCAGAGCAGCTCCCAAGTGTGT 1200  
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpPrAspSerLeuPro 420  
Db 1201 GAGACTCTGGAAGAGATCAGAGTTACTATACATCTCAGCATGGCCGGACAGCTGCCT 1260  
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
Db 1261 GACCTCAGCGTCTTCCAGAACCCTCAAGTAAATCCGGGAGCAATTTCTGCACATGGCGCC 1320  
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
Db 1321 TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGGCTGGGCTGCTCACTGAGGGAA 1380  
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480  
Db 1381 CTGGCAGTGGACTGGCCCTCATCCACCAATAACACCACCTCTGCTTCTGCACACAGTG 1440  
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
Db 1441 CCCTGGGACCAAGCTCTTTCCGAACCCGACCAAGCTCTGTCCACACTGCCAACCGCCA 1500  
Qy 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
Db 1501 GAGGACAGTGTGTGGCGAGGGCTGGCCCTGCCACAGCTGCGCGCCGAGGCACTGC 1560  
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
Db 1561 TGGGGTCCAGGGCCCAACCCAGTGTCAACTGCAGCCAGTTCTTCTCGGGGCCAGGAGTGC 1620  
Qy 541 ValGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
Db 1621 GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680  
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
Db 1681 TTGCCGTGCCACCTGAGTGTACGCCCCAGAAAGGCTCAGTGACCTGTTTGGACCGAG 1740  
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
Db 1741 GCTGACCAGTGTGTGGCTGTGCCCACTATAAGGACCCCTCCCTTCTCGSTGGCCCGCTGC 1800  
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
Db 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTGTTCAGATGAGAG 1860  
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640  
Db 1861 GGGCATGCGCAGCTTGGCCCATCACTGCACCCACCTCTGTGTGGACCTGGATGACAAG 1920  
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
Db 1921 GGCTGCCCGCCGAGCAGAGCAGCCAGCCCTCTGACGTCC 1959

RESULT 2

US-09-811-123-8



```

Db      1681  TTGGCGTCCACCTGAGTGTACGCCAGAAATGGCTCAGTACCTGTTTGGACCGGAG 1740
Qy      581  AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaAArgCys 600
Db      1741  GCTGACCAGTGTGGCGCTGTGCCCATATPAAGGACCTCCCTTCTCGTGGGCGCCGTGC 1800
Qy      601  ProSerGlyValLysProAspLeuSerTyrMetProLleTyrLysPheProAspGlu 620
Db      1801  CCCAGCGGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG 1860
Qy      621  GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db      1861  GGCGCATGCCAGCCTTGCCCATCACTGACCCACCTCTGTGTGGACCTGGATGACAAG 1920
Qy      641  GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db      1921  GGCTGCCCGCGGACGAGAGACCGCCCTCTGAGTCC 1959

RESULT 3
; Sequence 2, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-115-2

Alignment Scores:
Pred. No.:      0      Length:      3768
Score:          3628.00  Matches:      653
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:      9      Indels:      0
Db:              9      Gaps:      0

SEQ3 (1-653) x US-09-811-115-2 (1-3768)
Qy      1  MetGluLeuAlaAlaLeuCysArgTrrpGlyLeuLeuAlaLeuLeuProGlyVala 20
Db      1  ATGAGCTGGCGGCTGTGGCGCTGGGGGCTCTCTCGCCCTTTGGCCCCGGAGCC 60
Qy      21  AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db      61  GCGAGCACCAAGTGTGCACCGGCACACATGAAGCTGGCGCTCCCTGCCAGTCCCGAG 120
Qy      41  ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db      121  ACCACCTGGACATGCTCCCGCACCTTACAGGGCTGCCAGGTGGTGCAGGGAACCTG 180
Qy      61  GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db      181  GAACCTACCTACCTGCCCAATGCCAGCTGTCTCTCTGTCAGGATATCCAGGAGGTG 240
Qy      81  GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      241  CAGGGCTACGTGCTCATCGCTCACAACCAAGTAGGAGGCTGCCACTGCAGAGCTCGG 300
Qy      101  IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120

```

```

Db      301  ATTTGCGAGGACCCAGCTCTTTGAGGACAACATATGCCCTGGCGGTCTAGACAATGGA 360
Qy      121  AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db      361  GACCCGCTGAACAATACACCCCTGTACAGGGGCTTCCCAGAGAGGCTTGGGGAGCTG 420
Qy      141  GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db      421  CAGCTTCAAGCCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCAG 480
Qy      161  LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
Db      481  CTCTGCTACCCAGGACAGATTTGTGAAGACATCTTCCACAAGAACAACACCTGGCT 540
Qy      181  LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      541  CTCACACTGATAGACACCAACCGCTCTCGGGCTTGCACCCCTGTCTCCGATGTGAAG 600
Qy      201  GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      601  GGCTCCCGCTGTGGGAGAGAGTTCTGAGATTGTCAGAGCTTGACGCGACTGTCTGT 660
Qy      221  AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db      661  GCGGTGGCTGTGCCCTGCAAGGGCCACTGCCCTGCTGACTGCTGCTGCTGCTGCTG 720
Qy      241  AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db      721  GCTGCCGCTGCAGGGGCCCAACACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Qy      261  SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db      781  AGTGGCATCTGTGAGCTGCCTGCCAGCCCTGCTACCTACAAACACAGACACCTTTGAG 840
Qy      281  SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db      841  TCCATGCCCAATCCCGAGGGCCGATATACATTCGCGGCCAGCTGTGTGCTGCTGCTG 900
Qy      301  TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAspGln 320
Db      901  TACAACTACCTTCTACGGAGCTGGGATCTGCACCTCTGCTGCTGCTGCTGCTGCTGCT 960
Qy      321  GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db      961  GAGGTGACAGCAGAGGATGGAACACACGCGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1020
Qy      341  ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db      1021  GTGTGCTATGCTTGGGCATGGAGCACTTGCAGAGGTGAGGGCAGTTACCAAGTCCCAAT 1080
Qy      361  IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db      1081  ATCCAGAGGTTTGTGCTGCTCAAGAGATCTTTGGGAGCTTGGCATTTCTGCCGGAGAGC 1140
Qy      381  PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db      1141  TTTGATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCAGAGAGTCCCAAGTGTGT 1200
Qy      401  GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db      1201  GAGACTTGAAGAGATCAGAGGTTACCTATACATCTCAGCATGGCCGGACAGCTGCCT 1260
Qy      421  AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db      1261  GACCTCAGCTCTTCCAGAACCTTCAAGTAATCCGGGGAGCAATTTCTGCACATGGCGCC 1320
Qy      441  TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db      1321  TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGAA 1380
Qy      461  LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db      1381  CTGGGCAGTGGACTGGCCCTCATCCACCAACACCACCTCTGCTTCTGTCACACGCTG 1440

```

```
QY 481 ProTTPAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500
QY 501 GluAspGluCysValGlyGluGlnLeuAlaCysHisGlnLeuLeuAlaArgGlyHisCys 520
Db 1501 GAGGACGAGTGTGTGGGAGGCGCTGGCTGCCACCAAGCTGTGGCGCCGAGGCACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGTTCAGGCGCCACCCAGTGTGTCAACTCAGCCAGTCTCTTCGGGGGCCAGGAGTGC 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 GTGGAGGAATGCCGAGTACTCAGGGGCTCCCGAGGAGTAGTGAATGCCAGGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCGCTGCCACCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1741 GCTGACCAAGTGTGTGGCTGTGCCCACTATAAGGACCTCTCTCTGTGGCGCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTCTCCAGATCAGGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 1861 GCGCATGCCAGCCTTGCCCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 1920
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCGCAGAGAGCCAGCCCTCTGACGTCC 1959

RESULT 4
US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc-feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc-feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc-feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
```

```
; NAME/KEY: misc-feature
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

Alignment Scores:
Pred. No.: 0 Length: 3768
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservatives: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

SEQ3 (1-653) x US-09-854-356-9 (1-3768)
QY 1 MetGluLeuAlaLeuLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGGAGCTGGGGCGCTTGTGGCGCTGGGGCTCTCTCGCCCTCTTGCCTCCCGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGACACCCCAAGTGTGCACCGCACACATGAAGCTCGGCTCCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCACCTGGACATGCTCCGCCACCTCTACACAGGCTGCCAGGTGTGCAGGGAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 181 GAATCATCACTACCTGCCCAATGCCAGCTCTCTCTCGAGGATATCCAGGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTAGTGTCTATCGCTCACAAACAAAGTGAAGGAGGCTCCCACTGCAGAGGTGCGG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTTGGCGAGCACCAGCTCTTTGAGGACAACTATGCCCTGGCGTGTAGACAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCGCTGAACAATACCACCCCTGTACAGGGGCTCTCCCAAGGAGGCTTGGCGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTTCAGAGCTCACAGAGATCTTGAAGAGGAGGGTCTGTATCCAGCGGAACCCCGAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCTGCTACCAAGGACAGCATTTTGTGAAGGAGACATCTTCCACAAGAACAACAGCTGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTCTCCACCCCTGTCTCCGATGTGTAAAG 600
QY 201 GlySerArgCysTrpGlyLysSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGCTCTGGGGAGAGAGTCTCAGAGGATGTGTAGAGCTTCCAGCGGCACTGTCTGT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCCGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGTGCTGCATGAGCAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTCGGGCTGCAGGGGCCCAAGACACTCTGACTGCCTGGCTGCTCCACTTCAACCCAC 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGCATCTGTGAGCTGCACTGCCCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
```

Db	841	TCCATGCCCCAATCCCGAGGCGCGGTATACATTGGCGCCAGCTGTGTGACTGCTGTGCC	900
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACTACCTTCTACGAGCGTGGGATCTGCACCCCTCGCTGCCCCCTGCACACCAA	960
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerTysProCysAlaArg	340
Db	961	GAGGTGACAGCAGAGGATGGAACACACACGCGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1020
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValargAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGGTCTGGGCATGTGGAGCACTTCGAGAGGTGAGGCAGTTACCACTAGTCCAAT	1080
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGAGAGTTGCTGGCTGCAGAGAATCTTTGGAGCCTGGCATTTTCGCCGGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1141	TTTGATGGGACCCAGCCTCAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1200
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTGGAAGAGATCACAGGTTACTATACATCTCAGCATGGCGGGACAGCCCTGCCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCGTCTCCAGAACCTCGCAAGTAATCCGGGAGCAATCTGCACAATGGCGCC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCCTGCAAGGCTGGGCATCAGCTGGCTGGGGCTGGCTCCTCCTGAGGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGGCACTGGACTGGCCCTCATCCACATTAACACCCACCCTCTGCTGTGTGCACACGGTG	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCACTCTTTCGGAACCCGCACCAAGCTCTGTCTCCACACTGTCACACCGGCCA	1500
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACAGTGTGTGGCGAGGGCTGGCCCTGCCACCAGCTGTGGCGCCGAGGGCACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGGCCACCCAGTGTGTCAACTGCAGGCAGTTCCTTCTGGGGCCAGGAGTGC	1620
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValasnAlaArgHisCys	560
Db	1621	GTGGAGAAATGCCAGTACTGCAAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCTCTGAGTGTGACGCCCAAGTGGCTCAGTGACCTGTTTTGGACCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCAGTGTGTGGCTGTGCCACTATAAGGACCTTCCCTCTCTCGGTGGCCCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCAGCGGGTGTGAACCTGACCTCTCTACATCCCATCTGGAAGTTTCCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGCATGCCAGCTTGGCCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653

```
Db      481  |CTCTGCTACAGGACACGATT|TTGTGGAAGGACATCT|TCCACAAGAACACCAGTGCT| 540
Qy      181  |LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys| 200
Db      541  |CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTAAAG| 600
Qy      201  |GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys| 220
Db      601  |GGCTCCCGCTGCTGGGGAGAGATTCTGAGGATTGTCAGAGCTTGACGGCAGCTGCTGT| 660
Qy      221  |AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys| 240
Db      661  |CCCGTGCTGTGGCCGTGCAAGGGGCCACTGCCCACTGACTGCTGCATGAGCAGTGT| 720
Qy      241  |AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis| 260
Db      721  |GCTGCGGCTGCACGGGCCCAANGACACTCTGACTGCCCTGGCTGCCTCCACTTCAACCAC| 780
Qy      261  |SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu| 280
Db      781  |AGTGCATCTGTGAGCTGCACTGCCAGCCCTGGTGCTACCTACAACACAGACACGTTT| 840
Qy      281  |SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro| 300
Db      841  |TCCATGCCCAATCCCGAGGGCGGTATACATTGCGGCGCAGCTGTGTGACTGCCTGTCCC| 900
Qy      301  |TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln| 320
Db      901  |TACAACTACCTTTCTACCGAGCGTGGATCTCGACCCCTGCTGCGCCCTGCACAACCAA| 960
Qy      321  |GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg| 340
Db      961  |GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGACAGAACCCCTGTGCCCGA| 1020
Qy      341  |ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn| 360
Db      1021  |GTGTGCTATGCTGGCGATGAGCAGCTTGGCAGAGGTGAGGCGAGTTACCAGTGCCCAAT| 1080
Qy      361  |IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer| 380
Db      1081  |ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGGAGCCTGGCAATTTCTGCCGGAGAGC| 1140
Qy      381  |PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe| 400
Db      1141  |TTTGATGGGACCCAGCCCTCCAACTGCCCGCTCCAGCCAGACAGCTCCAAGTGT| 1200
Qy      401  |GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro| 420
Db      1201  |GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGGCGGACAGCGCTGCT| 1260
Qy      421  |AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla| 440
Db      1261  |GACCTCAGCGCTCTCCAGAAGCTGCAAGTAAATCGGGGACGAATTCGCAAAATGGCGCC| 1320
Qy      441  |TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu| 460
Db      1321  |TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGGAA| 1380
Qy      461  |LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal| 480
Db      1381  |CTGGCAGTGAGCTGGCCCTCATCCACCATAACACCCACCTCTGCTTGTGCACACGGTG| 1440
Qy      481  |ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro| 500
Db      1441  |CCCTGGGACCAAGCTCTTTGGAACCGCACCAGCTCTGCTCCACACTGCCCAACCGGCCA| 1500
Qy      501  |GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys| 520
Db      1501  |GAGGACGAGTGTGTGGCGAGGGCTGGCCCTGCACACAGCTGTGGCCCGGAGGGACATGC| 1560
Qy      521  |TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys| 540
```

```
Db      1561  |TGGGTCCAGGGCCCAACCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC| 1620
Qy      541  |ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys| 560
Db      1621  |GTGGAGGAATGCCAGTACTCAGGGGCTCCCGAGGAGTATGTGAAATGCCAGGCACTGT| 1680
Qy      561  |LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu| 580
Db      1681  |TTGCGTGCACCCCTGAGTGTGAGCCCAAGATGCTCAGTGACCTGTTTGGACCGGAG| 1740
Qy      581  |AlaaspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys| 600
Db      1741  |GCTGACCACTGTGGCTGTGCCCACTATAGGACCTCCCTTCTGCTGGCCCGCTGC| 1800
Qy      601  |ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProaspGluGlu| 620
Db      1801  |CCGAGCGGTGGAACCTGACCTCTCTACATGCCCATCTGGAAAGTTTCCAGATGAGGAG| 1860
Qy      621  |GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys| 640
Db      1861  |GGCGCATGCCAGCCTTGCCCATCACTGCACCCACTCCTGTGTGACCTGGATGACAAG| 1920
Qy      641  |GlyCysProAlaGluGlnArgAlaSerProLeuThrSer| 653
Db      1921  |GGCTGCCCGCGGAGCAGAGAGCCAGCCCTCTGACGTCC| 1959

RESULT 6
US-10-313-644-1
; Sequence 1, Application US/10313644
; Publication No. US20030157119A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: Hand-Zimmerman, Susan
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
; FILE REFERENCE: AND VIRUS-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.483C3
; CURRENT APPLICATION NUMBER: US/10/313,644
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-313-644-1

Alignment Scores:
Pred. No.: 0 Length: 3768
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ3 (1-653) x US-10-313-644-1 (1-3768)
Qy      1  |MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla| 20
Db      1  |ATGAGAGTGGCGGCTTGTGCGCTGGGGGCTCTCTCTCGCCCTTGTGCCCGGAGCC| 60
Qy      21  |AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu| 40
Db      61  |GCGAGCACCACAAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG| 120
Qy      41  |ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu| 60
Db      121  |ACCCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGGTGCAGGAACTG| 180
Qy      61  |GluLeuThrTyrLeuProThrAsnAlaSerPheLeuGlnAspIleGlnGluVal| 80
```



```

Db      |||||||TTCACCTACCTGCCCCACCAATGCCAGCGTGCCTTCTTCGAGGATATCCAGAGGTG 240
Qy      |||||||TyrSerLeuThrLeuGlnGluAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      |||||||TACTCGCTGACCTTGCAGGGCTGGCATCAGCTGGCTGGGCTCGCGCTCTACTGAGGAA 1380
Qy      |||||||LeuGlySerGlyLeuAlaLeuLeuHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db      |||||||CTGGGAGTGGAGCTGGCCCTCATCCACATAAACACCCACCTCTCTCTCGTCCACACGGTG 1440
Qy      |||||||ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db      |||||||CCCTGGGACAGCTCTTTTCGGAACCCGACCACTCTGTCTCCACACTGCCAACCCGCCA 1500
Qy      |||||||GluAspGluCysValGlyGluLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db      |||||||GAGGACAGGTGTGTGGCGGAGGGCTGGCCCTGCCACAGCTGTGGCGCCGAGGCGACTGC 1560
Qy      |||||||TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db      |||||||TGGGTCCAGGGCCCCCAGTGTCAACTGCAGCCAGTTCTCTTCGGGGCCAGGAGTGC 1620
Qy      |||||||ValGluCysArgValLeuGlnGlyLeuProArgGluThrValAsnAlaArgHisCys 560
Db      |||||||GTGGAGGAATGCCAGTACTGCAGGGCTCCCCAGGAGTATGTGAATGCCAGGCACTGT 1680
Qy      |||||||LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db      |||||||TTGCCGTGCCACCTGAGTGTACAGCCCAAGTGGCTCAGTACCTGTTTGGACCCGAG 1740
Qy      |||||||AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db      |||||||GCTGACCAGTGTGGCTGTGCCACATATAAGGACCTCCCTTCTCGTGGCCCGCTGC 1800
Qy      |||||||ProSerGlyValLysProAspLeuSerTyrMetProIleThrLysPheProAspGluGlu 620
Db      |||||||CCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTCGAAGTGTTCAGATGAGGAG 1860
Qy      |||||||GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db      |||||||GGCGATGCCAGCTTGGCCCACTCACTGCACCCACTCTCTGTGGACCTGGATGACAA 1920
Qy      |||||||GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db      |||||||GGTGGCCCCCGAGCAGAGAGAGCCGCTCTGACGTCC 1959

RESULT 7
US-09-441-411-5
; Sequence 5, Application US/09441411
; Publication No. US2003008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441.411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-5

Alignment Scores:
Pred. No.: 0 Length: 4473
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

```

```

Db      |||||||GAATCCTACCTGCCCCACCAATGCCAGCGTGCCTTCTTCGAGGATATCCAGAGGTG 240
Qy      |||||||GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      |||||||CAGGCTACGTGCTCTGCATCAACCAAGTGAAGGAGGTGCCACCTGCAGAGCTGGG 300
Qy      |||||||IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      |||||||ATTGTGGGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTCTAGACAAATGA 360
Qy      |||||||AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db      |||||||GACCCGCTGAACAATACCCCTGTTCACAGGGCCCTCCCCAGAGGCGCTGCGGGAGCTG 420
Qy      |||||||GlnLeuArgSerLeuThrGluLeuLysGlyGlyValLeuLeuGlnArgAsnProGln 160
Db      |||||||CAGCTTGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCCAG 480
Qy      |||||||LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db      |||||||CTCTGCTACCAAGGACAGATTTTGTGAAGGACATCTTCACAAAGAAACAACACAGCTGG 540
Qy      |||||||LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      |||||||CTCACACTGATAGACACCAACCCGCTCTCGGGCCCTGCCACCCCTGTTCTCCGATGTGA 600
Qy      |||||||GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      |||||||GGCTCCCGCTGCTGGGAGAGAGTCTGAGGATGTTCAGAGCCCTGACGCGCACTGTCTGT 660
Qy      |||||||AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db      |||||||GCCGGTGGCTGTGCCCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 720
Qy      |||||||AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db      |||||||GCTCCCGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGCCCTCCACTTCAACAC 780
Qy      |||||||SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db      |||||||AGTGGCATCTGTGAGCTGCCTGCCAGCCCTGGTCACTTACAAACACAGACAGCTTTGAG 840
Qy      |||||||SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db      |||||||TCCATGCCCAATCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGCTGCC 900
Qy      |||||||TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db      |||||||TACAACCTACCTTCTACGGAGCTGGGATCTCGACCCCTGCTGTGCCCCCTGCACAAACAA 960
Qy      |||||||GluValThrAlaGluAspGlyThrGlnArgCysGluLysSerLysProCysAlaArg 340
Db      |||||||GAGGTGACAGCAGAGGATGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCA 1020
Qy      |||||||ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db      |||||||GTGTGCTATGCTTGGCATGAGGACACTTCGAGAGGTGAGGCGAGTTACCACTGGCAAT 1080
Qy      |||||||IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db      |||||||ATCCAGGAGTTGCTGGCTCAAGAGATCTTTGGGAGCTGGCATTTCTGCGCGAGAGC 1140
Qy      |||||||PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db      |||||||TTTGATGGGACCCAGCTCCAACTGCTCCCGCTCCAGCCAGCAGAGCTCCAAGTGT 1200
Qy      |||||||GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db      |||||||GAGACTTGAAGAGATCAAGGTACCTATACATCTCAGCATGGCGGACAGCCCTGCCT 1260
Qy      |||||||AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440

```



DB:	11	Gaps:	0
SEQ3 (1-653) x US-09-441-411-5 (1-4473)			
Qy	1	MetGluLeuAlaLaLeuCysArgTrrpGlyLeuLeuLeuAlaLeuLeuProGlyVala	20
Db	175	ATGGAGCTGGGGCTTGCCGCTGGGGCTCCTCCTCGGCCHTTTGGCCCCCGGAGCC	234
Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	235	CGGAGCACCAAGTGTCCACGGGCACAGCATGAAGCTGGGGCTCCTGCCAGTCCCAG	294
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrglnGlyCysGlnValValGlnGlyAsnLeu	60
Db	295	ACCACCTGGACATGCTCCGCCACTCTACCAAGGGCTGCCAGGTGGTCAGGGAACCTG	354
Qy	61	GluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	355	GAACCTACCTACCTGGCCCAATAGCCAGCCTGTCTTCCTGCAGGATATCCAGGAGGTG	414
Qy	81	GlnGlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	415	CAGSGCTACGTGCTCATCGCTCACACAAGTAGCGAGGTCCACACTGCAGAGCGTCGGG	474
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAspAsnGly	120
Db	475	ATTGTGGGAGCACCCAGCTCTTTGAGGACAACCTATGCCCTGGCGCTGTAGACAATGGA	534
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Db	535	GACCCGCTGAACAATACACCCCTGTCCAGGGGCTCCCGAGGAGCGCTGCGGAGACTG	594
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160
Db	595	CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTGTATCCAGCGGAACCCCGA	654
Qy	161	LeuCystTyrrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
Db	655	CTCTGCTACCGAGGACAGATTTTGTGAAGGACATCTTCCACAAGAACAACCACTGGCT	714
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	715	CTCACACTGATAGACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAG	774
Qy	201	GlySerArgCystTpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	775	GGCTCCCGCTGCTGGGAGAGAGTTCGAGAGATTGTCAGAGCCTGACGGCACHTCTGT	834
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluInCys	240
Db	835	GCCGGTGGCTGTGCCGCTCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCACTGT	894
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	895	GCTCCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGCCCTGCCCTTCAACCCAC	954
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThrPheGlu	280
Db	955	AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGTGTCACTTACAACACAGACACGTTTTGAG	1014
Qy	281	SerMetProAsnProGluGlyArgTyrrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	1015	TCCATGCCCAATCCGAGGGCCGTATACATTCGGCGCCAGCTGTGTGACTGCTGCTGCC	1074
Qy	301	TyrAsnTyrrLeuSerThrAspValIGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	1075	TACAACATACCTTTCTACGGAGCTGGGATCTGTGACACCCCTCGTCTGCCCTCTGCACAACCAA	1134
Qy	321	GluValThrAlaGluAspGlyThrGlnArcCysGluLysCysSerLysProCysAlaArg	340
Db	1135	GAGGTGACAGCAGAGATGGAACACACGGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA	1194
Qy	341	ValCystTyrrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360

Db	1195		GTGTGCTATGCTGTGGGCATGGAGCACTTGCAGAGGTGAGGGCAGTTACCATGTGCCAAT	1254
Qy	361		IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1255		ATCCAGGAGTTTGTGGCTCAAGAAGACTTTTGGGAGCCTGGGCATTTCTGCCGGAGAGC	1314
Qy	381		PheAspGlyAspProAlaISerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1315		TTTGTATGGGACCCAGAGCTCCAACACTGCCCGCTCCAGCCAGACAGCTCCAAAGTGTTT	1374
Qy	401		GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1375		GAGACTCTGGAGAGATCACAGGTTACTATACATCTCAGCATGGCCGGACAGCTGCCT	1434
Qy	421		AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1435		GACCTCAGCGCTTTCAGAAACCTGCAAGTAATCCGGGACGAAATCTGCAACATGGCGCC	1494
Qy	441		TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1495		TACTCGGTGACCTTCGAAGGCTGGGCATCAGCTGGCTGGGGCTGCCTCCTACTGAGGAA	1554
Qy	461		LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1555		CTGGGCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCTGTGCACACGGTG	1614
Qy	481		ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1615		CCCTGGGACAGCTTCTTCGGAACCCGACCAACAGCTCTGCTTCCACACTGCCAACCGGCCA	1674
Qy	501		GluAspGlyCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1675		GAGACGAGTGTGGGCGAGGGCTTGGCTGCCACAGCTGTGGCCCGCCAGGGCAGCTGC	1734
Qy	521		TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1735		TGGGTTCCAGGGCCACCAGTGTCAACTGCAGCCAGTTCTCTTCGGGGCCAGAGTGC	1794
Qy	541		ValGluGlyCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1795		GTGAGGAATGCCAGTACTGCAGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACCTGT	1854
Qy	561		LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1855		TTGCCGTGCCACCTGAGTGTACGCCCAAGTGGCTCAGTGACCTGTTTGTGACCCGGAG	1914
Qy	581		AlaAspGlnCysValAlaCysAlaHisTyrIleAspProProPheCysValAlaArgCys	600
Db	1915		GCTGACCAGTGTGGGCTGTGCCCACTATAGAAGACCTCCCTTCTCGCTGGCCCGCTGC	1974
Qy	601		ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1975		CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCAGATGAGGAG	2034
Qy	621		GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2035		GGCGCATGCCAGCCTTCCCCCATCACTGCACCCACCTCCTGTGTGGAGCTGGATGACAAG	2094
Qy	641		GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2095		GGTGTCCCGCCGAGCAGAGAGAGCCAGCCCTCTGACGTCC	2133

RESULT 8  
US-10-101-510-81  
; Sequence 81, Application US/10101510  
; Publication No. US20030148295A1  
; GENERAL INFORMATION:  
; APPLICANT: WAN, JACKSON  
; APPLICANT: WANG, YIXIN  
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE  
; FILE REFERENCE: 15117.0012  
; CURRENT APPLICATION NUMBER: US/10/101,510

; CURRENT FILING DATE: 2002-03-20  
; PRIOR APPLICATION NUMBER: 60/276,947  
; PRIOR FILING DATE: 2001-03-20  
; NUMBER OF SEQ ID NOS: 805  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 81  
; LENGTH: 4473  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-101-510-81

## Alignment Scores:

Pred. No.:	0	Length:	4473
Score:	3628.00	Matches:	653
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

## SEQ3 (1-653) x US-10-101-510-81 (1-4473)

Qy	1	MetGluLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla	20
Db	175	ATGGAGCTGGGGCTTGTGCCGCTGGGGCTCCTCTCGCCCTTGGCCCGGAGCC	234
Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	235	CGGAGCACCCAAAGTGTGCACGGCACACATGAAGCTGGGCTCCCTGCCAGTCCCGAG	294
Qy	41	ThrHisLeuaspMetLeuArgHisLeuTyrglnGlyCysGlnValValGlnGlnLeu	60
Db	295	ACCCACCTGGACATGCTCCGCCACCTCTTACCAGGGCTGCCAGGTGGTGCAGGGAACCTG	354
Qy	61	GluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	355	GAACCTACCTACCTGCCACCAATGCCAGCTGTCTCTTCCAGGATATCCAGGAGTG	414
Qy	81	GlnGlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	415	CAGGGCTACGTGCTCATCGCTCACAACTAGTGGAGGAGTCCCACTGCAGAGCTGCGG	474
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAsnGly	120
Db	475	ATTGTGGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGCTGCTAGACAATGGA	534
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Db	535	GACCCGTGAACATACCACTCCCTGTACAGGGGCTCCCAAGAGGCTCGCGGAGCTG	594
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160
Db	595	CAGCTTCGAAGCCTCAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCGAG	654
Qy	161	LeuCysTyrrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
Db	655	CTCTGCTACCAAGGACAGATTTTGTGAAGGACATCTTCCACAGAACCAACCACTGGCT	714
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	715	CTCACATGATAGACACCAACCGCTCTCGGGCTTCCACCCCTGTTCTCCGATGTGAAG	774
Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	775	GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGACAGCCCTGACGGCGACTGTCTGT	834
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys	240
Db	835	GCCGGTGGCTGTGCCCTGCAAGGGGCCACTGCCCACTGACTCTGCTGCATGACAGTGT	894
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	895	GCTGCCGCTGCACGGGCCCCAAGCACTCTCTGACTGCTGCTGGCTGCTCCCTTCAACCAC	954

Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThrPheGlu	280
Db	955	AGTGGCATCTGTAGCTGCACCTGCCAGCCCTGTGTACCTACCAACACAGACACGTTTGAG	1014
Qy	281	SerMetProAsnProGluGlyArgTyrrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	1015	TCCATGCCCAATCCGAGGCGCGGTATACATTCGGCGCCAGCTGTGTACTGCTGTGCTCC	1074
Qy	301	TyrAsnTyrrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	1075	TACAACTACCTTCTACGGAGGTGGATCTGTGCACCTCGTCTGCCCTTGCACCAACCA	1134
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	1135	GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1194
Qy	341	ValCysTyrrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1195	GTGTGCTATGTCTGGGCATGGAGCACTTGGAGAGGTGAGGGCAGTTACCAAGTGCCTAA	1254
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1255	ATCCAGAGTTTGTGCTGCAAGAAGATCTTTGGAGCCCTGGCATTTCTGCCGAGAGC	1314
Qy	381	PheaspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1315	TTTGATGGGACCCAGAGCTCCACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1374
Qy	401	GluThrLeuGluIleThrGlyTyrrLeuTyrrIleSerAlaTrpProAspSerLeuPro	420
Db	1375	GAGACTCTGAAGAGATCAGAGTTACCTATACATCTCAGCATGGCCGAGAGCTGCCT	1434
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1435	GACCTCAGCGCTTCCAGAACCTCCTCAAGTAATCCGGGAGCAATCTGCACAAATGGCC	1494
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1495	TACTCGTGCACCTTGCAGGGCTGGGATCAGCTGGCTGGGGCTGCCTCACTCAGGGAA	1554
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1555	CTGGGAGTGGACTGGCCCTCATCCACATAACACCCACCTCTGCTTCTGCACACCGTG	1614
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1615	CCCTGGGACCAAGCTCTTTGGGAACCCGCCAAGCTCTGTCTCCACTGCCAACCGGCCA	1674
Qy	501	GluaspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1675	GAGGACGAGTGTGTGGGCGAGGGCTGGCTGCCACCAAGCTGTGCGCCCGAGGCACTGC	1734
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1735	TGGGTCACAGGGCCACCCAGTGTCAACTGCAAGCCAGCTTCTTTCGGGGCCAGAGTGC	1794
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrrValAsnAlaArgHisCys	560
Db	1795	GTGGAGGAATGCCAGGACTTGCAGGGGCTCCCAAGGAGTATGTGAATGCCAGGCACTGT	1854
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1855	TTGCCGTGCCACCTGAGTGTGAGCCCGCAGAAATGGCTCAGTGACGTGTTTGGACCGAG	1914
Qy	581	AlaaspGlnCysValAlaCysAlaHisTyrrLysAspProProPheCysValAlaArgCys	600
Db	1915	GCTACCAAGTGTGGGCTGTGCCACTATTAAGACCCCTCCCTTCTGCGTGGCCGCTGC	1974
Qy	601	ProSerGlyValLysProAspLeuSerTyrrMetProIleTrpLysPheProAspGlu	620
Db	1975	CCCAGCGGTGGAACCTGACCTCTCTACATGCCCATCTCGAAGTTTCCAGATGAGGAG	2034
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640

```

Db 2035 GCGCATGCCAGCCTTGCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAG 2094
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 2095 GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGAGGTCC 2133

RESULT 9
US-10-146-473-32
; Sequence 32, Application US/10146473
; Publication No. US20030108888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-473-32

Alignment Scores:
Pred. No.: 0 Length: 4473
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

SEQ3 (1-653) x US-10-146-473-32 (1-4473)
QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 175 ATGGAGCTGGCGGCTGTGCGCTGGGGCTCTCCCTCGCCCTTGTGCCCGGAGCC 234
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCGAGCACCCCAAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 294
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 295 ACCACCTGGACATGCTCCCGCCTCTTACCAGGGCTGCCAGGTGGTCAGGGAAACCTG 354
QY 61 GluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 355 GAATCTACCTACCTGCCCCACCAATGCCAGCGCTGCTTCTCTCCAGGATATCCAGAGGTG 414
QY 81 GlnGlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGGCTACGTGCTCATGCTGCACAAACCAAGTAGGAGGAGTCCCACTGCAGAGGTGGG 474
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAspAsnGly 120
Db 475 ATTGTGGAGGACCCACCTCTTTGAGGACAACTATGCCCCCTGGCCGCTAGACAAATGGA 534
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGCTGACATAACACCCCTGTACAGGGGCTCCCGAGGAGGCTCGGGAGGTG 594
QY 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 595 CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTTGATCCAGCGGAACCCCGAG 654

```

```

QY 161 LeuCysTyrrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 655 CTCGTCTACAGGACACGATTTTGTGGAAGNACATCTTCCACAGAACACACCGCTGGCT 714
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 715 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 774
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 775 GGCTCCCGCTGCTGGGGAGAGAGTTCTGAGGATTCAGAGGCTCAGCGCACTGCTGT 834
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 835 GCGGGTGGCTGTGCCCGCTGCAAGGGCCACTGCCCACTGACTGCTGCCCTGGCTTCAAC 894
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 895 GCTGCCGCTGCACGGGGCCCCAAGCACCTCTGACTGCTGCCCTGGCTTCAACAC 954
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThrPheGlu 280
Db 955 AGTGGCATCTGTGAGCTGACTGCCAGCCCTGGTGCCTACAAACACAGACAGCTTTGAG 1014
QY 281 SerMetProAsnProGluGlyArgTyrrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 1015 TCCATGCCCAATCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGACTGCTCCTGCC 1074
QY 301 TyrAsnTyrrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1075 TACAACCTACCTTTCTACGGAGCTGGATCCTGCACCCCTCGTGCCTGCCCAACCAA 1134
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1135 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAAGTGAGCAAGCCCTGTGCCGA 1194
QY 341 ValCysTyrrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1195 GTGTGCTATGCTTGGGCTGAGGACCTTGGAGAGGTGAGGGCAGTTACCAAGTCCCAAT 1254
QY 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1255 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGAGCCTTGGCATTTCTGCCGAGAGC 1314
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1315 TTTGATGGGGACCCAGCCCTCCAAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT 1374
QY 401 GluThrLeuGluGluIleThrGlyTyrrLeuTyrrIleSerAlaTrpProAspSerLeuPro 420
Db 1375 GAGACTCTGGAGAGATACACAGTTACCTATACATCTCAGCATGCCCGGACAGCCCTGGCT 1434
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1435 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGACGAATTTCTGCACAATGGCGCC 1494
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1495 TACTCGCTGACCCCTGCAAGGGCTGGCATCAGCTGGCTGGGGCTCGCTCCTACTGAGGAA 1554
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1555 CTGGCAGTGACGTGGCCCTCATCCACCATACACCCACCTCTGCTTCTGTCACACCGGTG 1614
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1615 CCCTGGGACCAAGCTCTTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCCAACCGGCA 1674
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1675 GAGGACAGTGTGTGGCGAGGGCTGGCTGCCACCAAGCTGTGCCCGCCGAGGCACTGTC 1734

```

```

QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1735 TGGGTCCAGGGCCACCAGTGTCTCACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1794
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1795 GTGGAGGAATCCGAGTACTCAGAGGGCTCCAGGGAGTATGTAATGCCAGGCACATGT 1854
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1855 TTGGCGTCCACCTTGAGTGTACGCCCCAGAAATGGCTCAGTGACCTGTTTGGACCCGAG 1914
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1915 GCTGACCAAGTGTGGCCCTGTGCCACTATTAAGACCTCCCTTCTGCTGGCCGCGTGC 1974
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleThrLysPheProAspGluGlu 620
Db 1975 CCCAGCGGTGTGAACCTTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG 2034
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 2035 GGCGATGCCAGCCTTGCCCATCAACTGCACCCACCTCCTGTGTGGACCTGGATGACAAG 2094
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 2095 GGCTGCCCGCCGAGCAGAGACCCAGCCCTCTGACGTCC 2133

RESULT 10
US-10-207-655-44
; Sequence 44, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Lebetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-207-655-44

Alignment Scores:
Pred. No.: 0 Length: 4473
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

SEQ3 (1-653) x US-10-207-655-44 (1-4473)
QY 1 MetLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db 175 ATGGAGTGGCGGCTTGTGCGCGTGGGGCTCTCTCTCGCCCTCTTGGCCCGGAGCC 234
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCAGACCCCAAGTGTGCACCCGACAGACATGAAGTGGGCTCCCTGCGAGTCCCGAG 294
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 295 ACCACCTGGACATGTCTCCGCCACCTCTACACAGGCTGCCAGTGGTGGAGGAAACCTG 354
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 355 GAACTCACCTTACCTGCCCAATGCCAGCTGTCTTCTGTCAGGATATCCAGGAGGTG 414

```

```

QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGCTACGTGCTCATCTGCTCACAAACCAAGTAGGAGGCTGCCACTGCAGAGGTGGCG 474
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValIleAspAsnGly 120
Db 475 ATTGTGCGAGCACCACCTCTTTGAGACAACTATGCCCTGGCGCTGTAGACAATGGA 534
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGTGAACAATACCAACCCCTGTCTACAGGGGCTCCCCAGGAGGCTGCGGGAGCTG 594
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 595 CAGCTTGCAAGCCCTCACAGATCTTGAAGAGGGGTCTTGTATCCACAGCAACCCAG 654
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
Db 655 CTCTGTACCAAGGACACGATTTTGGGAAGGACATCTTCCACAGAAACACGAGCTGCT 714
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 715 CTCACACTGATAGACACCAACCCCTCTCGGGGCTGCCACCCCTGTCTCCGATGTGTAAG 774
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 775 GGCTCCGCTGCTGGGAGAGAGTCTGAGGATTTGTACAGAGCTGTACAGCGCTGTCTGT 834
QY 221 AlaGlyLysCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 835 GCCGGTGTGTCGCCCTGCAAGGGCCACTGCCACTGACTGTGTCATGACGAGTGT 894
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 895 GCTCCGGCTGCACGGGCCCAACGACACTCTGACTGCTGCTGGCTGCCCTCCACTCAACCAC 954
QY 261 SerGlyLysCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 955 AGTGGCATCTGTGAGCTGCACGTGCCAGCCCTGCTACCTACAAACACAGACACCTTTGAG 1014
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 1015 TCCATGCCCAATCCCGAGGGCCGTATACATTCGGCGCCAGCTGTGTGACTGTCTGTCCC 1074
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1075 TACAACCTACCTTCTACGGAGTGGGATCTGTGACCCCTCTGCTGCCCTGCACAAACCA 1134
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1135 GAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGAAAGCCCTGTGCCGA 1194
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1195 GTGTGCTATGCTTGGCATGGAGCACTTGCAGAGGTGAGGGCAGTTACAGTGCACAT 1254
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1255 ATCCAGGAGTTTGTGCTGCAAGAAGATCTTTGGGAGCCCTGGCATTTCTGCCGAGAGC 1314
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe 400
Db 1315 TTTGATGGGACCCAGCCTCCAACTGCCCGCTCCAGCGCTCCAGCAGAGCTCCAAAGTCTTT 1374
QY 401 GluThrLeuGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1375 GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGGCGGAGAGCTGCCT 1434
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1435 GACCTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGGAGCAATCTGCAACATGGCGCC 1494
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460

```

```

Db      1495  TACTCGCTGACCTGCAAGGGCTGGGCATCAGTGGCTGGGCTGGCTCACTGAGGGAA 1554
QY      461  LeuGlySerGlyLeuAlaLeuLeuHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db      1555  CTGGGCGAGTGGAGTGGCCCTCATCCACCATACACCCACTCTGCTGCTGCACACGGTG 1614
QY      481  ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisHisThrAlaAsnArgPro 500
Db      1615  CCTGGGACCACTCTTCGGAAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1674
QY      501  GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db      1675  GAGGACGAGTGTGTGGCGAGGGCTGGCTGCCACCACTGTGCCCGCCAGGGCACTGC 1734
QY      521  TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db      1735  TGGGGTCCAGGGGCCCAACCCAGTGTGTCAACTGCAGCCAGTTCTTCTGGGGCCAGGAGTGC 1794
QY      541  ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db      1795  GTGGAGGAATGCCGAGTACTGCAGGGCTCCCGAGGAGTATGTGATGCCAGGCACTGT 1854
QY      561  LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db      1855  TTGCCGTGCCACCTGAGTGTGAGCCCAAGATGGCTCAGTGACCTGTTTGGACCGGAG 1914
QY      581  AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db      1915  GCTGACCACTGTGTGGCTGTGCCCACTATAGGACCCCTCCCTTCTGCGTGGCCGCTGC 1974
QY      601  ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db      1975  CCCAGGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTCCAGATGAGGAG 2034
QY      621  GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db      2035  GCGCGATGCGACCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAG 2094
QY      641  GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db      2095  GGCTGCCCCCGCCAGCAGAGAGCCAGCCCTCTGACGTCC 2133

```

RESULT 11

```

US-09-877-177-11
; Sequence 11, Application US/09877177
; Publication No. US20020192652A1
; GENERAL INFORMATION:
; APPLICANT: Peter V. Danenberg et al.
; TITLE OF INVENTION: Method of determining Epidermal Growth
; FILE OF INVENTION: Factor Receptor and HER2-New Gene Expression
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177-11

```

```

Alignment Scores:
Pred. No.: 0 Length: 4530
Score: 3628.00 Matches: 633
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

```

SEQ3 (1-653) x US-09-877-177-11 (1-4530)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20

```

Db      151  ATGGAGCTGGCGCCTTGTGGCGCTGGGSCCTCTCTCTCGCCCTCTTGGCCCCCGAGGCC 210
QY      21  AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db      211  GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 270
QY      41  ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db      271  ACCCAGCTGGACATGCTCCGCCACCTCTACAGGGCTGCAGGTGGTGCAGGAAACCTCG 330
QY      61  GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db      331  GAACTCACCTACCTGCTGCCACCAATGCCAGCCTGTCTCTCTGCAGGATATCCAGGAGGTG 390
QY      81  GlnGlyTyrValLeuIleAlaHisGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      391  CAGGGCTAGCTGCTCATCGCTCACACCAAGTGCAGCAGGTCCACCTGCAGAGGCTGCGG 450
QY      101  IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      451  ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCCTGCTAGACAATGGA 510
QY      121  AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db      511  GACCCGCTGAACAATACCACTGTCACAGGGGCTCCCGAGGAGGCTGCGGGAGCTG 570
QY      141  GlnLeuArgSerLeuThrGluLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db      571  CAGCTTCGAAGCTCACAGAGATCTTGAAGAGAGGGTCTTGATCCAGGCGAACCCTCCAG 630
QY      161  LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db      631  CTCTGCTACCAGACACGATTTTGTGAAGAGACATCTTCCACAAGAAACCAACAGCTGGCT 690
QY      181  LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      691  CTACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTGTCTCCGATGTGTAAG 750
QY      201  GlySerArgCysTrpGlyCysSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      751  GGCTCCCGCTGCTGGGGAGAGAGTTCTTGAGGATTTGTCAGAGCCTGCAGCGCACTGTCTGT 810
QY      221  AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db      811  GCGGTGGCTGTGCCCGTGCAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 870
QY      241  AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db      871  GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGGCTGCCCTCCACTTCAACCA 930
QY      261  SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db      931  AGTGGCACTGTGAGCTGCACCTGCCAGCCCTGGTGCACCTACAACACACAGACAGTTGAG 990
QY      281  SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db      991  TCCATGCCCAATCCCGAGGGCCGTTATACATTCGGGGCCAGCTGTGTGACTGCTGCTCC 1050
QY      301  TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db      1051  TACAACCTACCTTCTACGGAGCTGGGATCCTGCACCTGCTGCGCCCTGCACCAACCAA 1110
QY      321  GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db      1111  GAGGTGCACAGCAGGATGGAAACAGCGGTGTGAGAAGTGCAGAAAGCCCTGTGCCCGA 1170
QY      341  ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db      1171  GTGTGCTATGGTCTGGGCATGGAGCACCTTCCAGAGGTGAGGGCAGTTACCACTGCCAAT 1230
QY      361  IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380

```

```

Db 1231 ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGGATTTCTCCCGGAGAGC 1290
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1291 TTTGATGGGACCCAGAGCTCCACACACTCCCGCTCCAGCCAGAGCAGCTCCAGTGT 1350
Qy 401 GluThrLeuGluGluLeuThrGlyTyrLeuTyrIleSerAlaTrpAspSerLeuPro 420
Db 1351 GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGCGCGAGCAGCTGCT 1410
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1411 GACCTCAGCGTCTCCAGAACCTCAAGTAATCCGGGAGCAATCTGCAACATGGCGCC 1470
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1471 TACTCGTGACCTTGCAGGGCTGGCATCAGCTGGCTGGGGCTCGGCTCACTGAGGGA 1530
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1531 CTGGGAGTGAGCTGGGCTCATCCACCAACACCCACCTCTGCTTGTGCACACGGTG 1590
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1591 CCCTGGGACCACTCTTTCGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCA 1650
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1651 GAGGACGAGTGTGGGAGGGCTGGCTGCCACCAAGCTGTCGGCCGAGGAGCACTGC 1710
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1711 TGGGTCCTCAGGGCCACCCAGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGAGTGC 1770
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1771 GTGAGGAAATCCAGAGTACTCAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1830
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1831 TTGCGTGCCACCTGTAGTGTGCAGCCCGAGAAATGGCTCAGTGACCTGTTTGGACCGAG 1890
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrIlyAspProProPheCysValAlaArgCys 600
Db 1891 GCTGACCAAGTGTGGCTGTGGCCACTATAGGACCCCTCCCTTCTGCGTGGCCGCTGC 1950
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1951 CCCAGCGGTGTAAACCTGACCTCTCTACATGCCCACTATGGAAGTTTCCAGATGAGAG 2010
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 2011 GCGCATGCCAGCTTGCCTCATCACTGACCCACCTCTCTGTGGACCTGGATGACAAG 2070
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 2071 GGCTGCCCGCCGAGAGAGAGCCAGCCCTCTGACGCTCC 2109

```

RESULT 12

```

; US-10-007-926A-119
; Sequence 119, Application US/10007926A
; Publication No. US20030143539A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007,926A

```

```

; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 119
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: v-erb-b2 avian erythroblastic leukemia viral
; OTHER INFORMATION: oncogene homolog 2 (neuro/glioblastoma derived
; OTHER INFORMATION: oncogene homolog) (ERBB2) gene.
US-10-007-926A-119

```

Alignment Scores:

```

Pred. No.: 0 Length: 4530
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

```

SEQ3 (1-653) x US-10-007-926A-119 (1-4530)

```

Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGGAGCTGGGGCTGTGGCGCTGGGGCTCTCTCTCGCCCTTTCGCCCGCGGAGCC 210
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 GCGAGCACCAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 270
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValGlnGlyAsnLeu 60
Db 271 ACCACCTGGACATGTCTCCCGCCACTCTTACCAGGGCTGCCAGTGGTGCAGGAAACCTG 330
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 331 GAACCTACCTACCTGCCACCAATGCCAGCTGCTCTCTCTGCAGGATATCCAGGAGTG 390
Qy 81 GlnGlyTyrValLeuLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTACGTGCTCATCGCTCACACCAAGTGCAGGAGGTCCTCCCTGCAGGAGTCCGG 450
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGCGAGCACCAGCTCTTTGAGGACAACTATGCTCCCTGGCGCTGCTAGACAATGA 510
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGCTGAACAATACCACTCCAGGGCTCCAGGAGGCTCCCGAGGAGGCTGCGGAGCTG 570
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 571 CAGCTTCAAGCCCTCACAGAGATCTTGAAGAGAGGGGTCTTGATCCAGCGGAAACCCCGAG 630
Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGCTACCAAGGACAGATTTTGTGAAGAGACATCTTCCACAAGACAACACCTGGCT 690
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACACCAACCGCTCTCGGSCCTGCCACCCCTCTTCTCCGATGTGTAAG 750
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCCGCTGCTGGGAGAGAGTCTCAGGATTTCTCAGAGCTTGTGAGCGGACCTGCTGT 810
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db 811 GCGGCTGGCTGTGCCCTGCAAGGGCCACTGCCACTGACTGCTGCTGCTGCTGCTGCTGCT 870

```

```
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 871 GCTGCGGCTGCACGGGCCCAAGCACACTCTGACATGCGCTGCGCTCCACTTCAACCCAC 930
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 931 AGTGGCACTGTGAGCTGCATGCCACCGCTGTGTCACCTTCAACACACAGACACGTTTGAG 990
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 991 TCCATGCCAATCCGAGGGCGGTATACATTGGCGCCAGCTGTGTGACTGCTGTGCTCC 1050
QY 301 TyrAsnTyrLeuSerThrAspValCysSerCysThrLeuValCysProLeuHisAsnGln 320
Db 1051 TACAACCTACCTTCTACGACGCTGGGATCTCTGCACCTCGTGTGCGCCCTGCACAACCAA 1110
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1111 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAAGTGCAGCAAGCCCTGTGCCGA 1170
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1171 GTGTGCTATGTGTGGGATGGACACTTGGCAGAGGTGAGGCGAGTTTACCAGTGCCTAAT 1230
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1231 ATCCAGGAGTTGTGCTGCGCAAGAAGATCTTTGGGAGCCTGGCATTCTGCCGGAGAGC 1290
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
Db 1291 TTTGATGGGACCCAGCCCTCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT 1350
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420
Db 1351 GAGACTCTGGAAGAGATCACAGGTACCTATATCATCTCAGCATGGCCGGCAGCCCTGCCT 1410
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1411 GACCTCAGCGCTTCCACAACCTGCAAGTAATCCGGGACGAAATCTGCACAATGGCGCC 1470
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1471 TACTTGCTGACCCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTACTGAGGAA 1530
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1531 CTGGCAGTGAGCTGGCCCTCATCCACCATACACCCACCTCTGCTTCTGTCACACGGTG 1590
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1591 CCCTGGGACCACTCTTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCAGGCA 1650
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgLysHisCys 520
Db 1651 GAGGACGAGTGTGTGGCGAGGGCTGGCCCTGCCACCACTGTCGCCCGCAGGCGACATGC 1710
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1711 TGGGTCCAGGGCCCCCAGGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1770
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1771 GTGGAGGAATGCCGAGTACTCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACATGT 1830
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1831 TTGCGTCCCACTTGTGAGTGTGAGCCCGCAGCAATGGCTCAGTGACCTGTTTGGACCGGAG 1890
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPropheCysValAlaArgCys 600
Db 1891 GCTGACCAAGTGTGGCGCTGTGCCACTATAAGGACCCCTTCTGCTGGCGCGCGCTGC 1950
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
```

```
Db 1951 CCCAGCGTGTGAACCTTGACCTCTCCATGATGCCCATCTGGAAGCTTTCAGATGAGGAG 2010
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 2011 GGCGCATGCCAGCCTTGCCCATCAACTGACCCACTCTCTGTGTGACCTGGATGACAAG 2070
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 2071 GGCTGCCCGCGGACGACAGAGCCAGCCCTCTGACGTCC 2109

RESULT 13
US-10-101-510-124
; Sequence 124, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; PRIORITY FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 124
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-124

Alignment Scores:
Pred. No.: 0 Length: 4530
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

SEQ3 (1-653) x US-10-101-510-124 (1-4530)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGAGCTGGCGCCTTGTGCGCTGGGGCTCTCTCGCCCTCTTGCCCCCGGAGCC 210
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 GCGAGCACCAAGTGTGACCGGCGACAGACATGAAGCTGCGGCTCCCTGCCAGTCCGAG 270
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAACCTCG 330
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 331 GAATCAGCTACCTGCCCCACCAATGCCAGCCTGCTCTCTCCAGGATATCCAGAGGTG 390
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGCTACGTGCTCATGCTCACAAACCAAGTGAGGACAGGTCCCACTGCAGAGGTGCGG 450
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTTGGCAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGCTGTAGACAATGGA 510
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGCTGAACATATACCACCCCTGTACAGGGGGCTCCCGAGGAGGCTGCGGGAGCTG 570
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 571 CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCCAG 630
```



```

QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGCTACCAAGACACGATTGTTGGAAAGACATCTTCCACAAGAACACAGCTGGCT 690
QY 181 LeuThrIleuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACATGATAGACACCAACCGCTCTCGGGCCCTGCCACCCCTGTTCTCCGATGTGAAG 750
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCCTCCGCTCTGGGAGAGAGTCTGAGAGATTGTGAGACCTTGACGGCAGCTGTCTGT 810
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db 811 GCCGGTGGCTGTGCCCGCTGCAAGGGCCACTGCGCCACTGACTGCTGCCATGACGAGTGT 870
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 871 GCTGCCGCTGCACGGGCCCAAGCACTCTGACTGCGCTGGCGCTGCCCTCCACTTCAACAC 930
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 931 AGTGGCATCTGTAGCTGCACTGCCAGCCCTGGTCACTACCAACACAGACACGTTTGAG 990
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 991 TCCATGCCCAATCCGAGGGCGGTATACATTCCGGCCAGCTGTGTGACTGCGCTGTGCC 1050
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1051 TACAACTACCTTTCTACGGAGCTGGGATCTCTGCACCCCTGCTGCTGCCCTGCACAAAC 1110
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1111 GAGGTGACAGCAGAGGATGAACACACGCGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1170
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1171 GTGTGCTATGCTTGGGCATGGAGCACTTCCGAGAGTGGGCGAGTTACCAAGTGCCT 1230
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1231 ATCCAGGAGTGTCTGCTGCACAGAGATCTTTGGGAGCTGGCATTTCTGCCGGAGC 1290
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
Db 1291 TTTGATGGGACCCAGCCTCCAACTGCCCCGCTCCAGCCAGCAGCAGCTCCAAGTGT 1350
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1351 GAGACTTGGAGAGATCACAGGTTACCTATACATCTCAGCATGGCGGACAGCCTGCCT 1410
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyVala 440
Db 1411 GACCTCAGCGCTCTCCAGAACCTGCAGTAATCCGGGGAGCAATCTGCACATGGCGCC 1470
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1471 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGAA 1530
QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1531 CTGGCAGTGGATGGCCCTTCATCCACCAATTAACACCCACTCTGCTTCTGTCACAGGGT 1590
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1591 CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACTGCGCAACCGGCCA 1650
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1651 GAGGACGAGTGTGTGGCGCAGGGCCCTGGCCTGCCACCACTGCTGCGCCCGGAGGCACTGC 1710

```

---

```

QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1711 TGGGTCAGGGCCACCCAGCTGTCAACTGCAGCAGTTCCTTCGGGGCCAGAGTGC 1770
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1771 GTGGAGGAATCCGAGTACTGCAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACGT 1830
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1831 TTGCGGTGCCACCTGAGTGTGAGCCAGAAATGGCTCAGTACCTGTGTTGGACCGGAG 1890
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys 600
Db 1891 GCTCACCAGTGTGTGGCTGTGCCACTATATAAGACCCCTCCTCTCTCGTGGCCGCTGC 1950
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1951 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 2011 GGCGCATGCCAGCTTCCCCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 2070
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 2071 GGCTGCCCGCCGAGCAGAGAGAGCCAGCCCTCTGACGTCC 2109

```

RESULT 14

US-10-338-730-1

; Sequence 1, Application US/10338730

; Publication No. US20030147905A1

; GENERAL INFORMATION:

; APPLICANT: Genzyme Corporation

; APPLICANT: Nicolette, Charles A.

; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS

; FILE REFERENCE: 5017C

; CURRENT APPLICATION NUMBER: US/10/338,730

; CURRENT FILING DATE: 2003-01-08

; PRIOR APPLICATION NUMBER: US 09/527,487

; PRIOR FILING DATE: 2002-03-16

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 4530

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (151)..(3915)

; OTHER INFORMATION:

US-10-338-730-1

Alignment Scores:

Pred. No.: 0 Length: 4530

Score: 3628.00 Matches: 653

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 12 Gaps: 0

SEQ3 (1-653) x US-10-338-730-1 (1-4530)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyVala 20

Db 151 ATGAGACTGGCGGCTTGTGCCCTGGGGCTCCTCTCTGCCCTCTTGGCCCCCGGAGCC 210

QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40

Db 211 GCAGACCAACCAAGTGTGCACCGCAGACATCAAGCTGGGGCTCCCTGCTCCAGTCCCGAG 270

QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60



Db 271 ACCACCTGGACATGCTCGCCACCTCTTACCAGGGCTGCCAGTGGTGCAGGGAACCTG 330  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
 Db 331 GAACCTACCTACCTGCCCCACCAATGCCAGCCTGTCTTCTTCGAGGATATCCAGGAGTG 390  
 QY 81 GlnGlyTyrValLeuIleAlaHisGlnValArgGlnValProLeuGlnArgLeuArg 100  
 Db 391 CAGGGCTACGTGCTCATGCTCACAACCAAGTGAGCAGGTCCCACTGCAGAGGCTGCGG 450  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 Db 451 ATTGCGAGGACCCAGCTCTTTGAGGACACTATGCTTGGCCGTGCTAGACATGGA 510  
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 Db 511 GACCCGCTGAACAATACCACCTCTGCACAGGGGCTCCCCAGGAGGCTGCGGAGCTG 570  
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160  
 Db 571 CAGCTTCGAAGCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCACG 630  
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180  
 Db 631 CTCTGCTACCGAGACAGATTTTGTGAAGGAGCATCTTCCCAAGAACCAACACGCTGGCT 690  
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 Db 691 CTCACACTGATAGACCAACACGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTAG 750  
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 Db 751 GGCTCCCGCTGTGGGGAGAGAGTCTGAGGATGTGTACAGGCTGACGGGCATGCTCTGT 810  
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240  
 Db 811 GCCGGTGGCTGTGCCCGCTGCAAGGGGGCCACTGCCCACTGACTGTGCTGCCATGAGCAGTGT 870  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 Db 871 GCTGCCGGCTGACGGGCCCCAAGCAGCTCTGACTGGCTGGCTGCTTCCACTTCAACCCAC 930  
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 Db 931 AGTGGCATCTGTGAGCTGACCTGCCAGCCCTGGTCACCTCAACACACAGACAGCTTTGAG 990  
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 Db 991 TCCATGCCCAATCCCGAGGGCGGTATACATTCCGGCGCCAGCTGTGTGACTGCCCTGTCCC 1050  
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 Db 1051 TACAACCTACCTTTCTACGGAGCTGGGATCTCGACCTCGTCTGCCCTCGCCCTGCACACCA 1110  
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 Db 1111 GAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1170  
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 Db 1171 GTGTGCTATGTGCTGGCATGAGGACCTTGCAGAGGTTGAGGCGAGTTACCAAGTCCCAAT 1230  
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 Db 1231 ATCCAGGAGTTTGTCTGGCTGCAAGAAGATCTTTGGGAGCTTGGCATTTCTGCGGAGAGC 1290  
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
 Db 1291 TTTGATGGGACCCAGCTCCCAACACTGCCCGCTCCAGCCAGACAGCTCCCAAGTGT 1350  
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro 420  
 Db 1351 GAGACTCTGGAAGAGATCACAGTTTACCTATACATCTCAGCATCTCAGCATGCGCGACAGCTGCT 1410

RESULT 15

US-10-177-293-125

; Sequence 125, Application US/10177293

; Publication No. US20030124128A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Glatt, Karen

; APPLICANT: Zhao, Xumei

; APPLICANT: Gannavarpu, Manjula

; APPLICANT: Kamatkar, Shubhangi

; APPLICANT: Mertens, Maureen

; APPLICANT: Myer, Vic

; APPLICANT: Wang, Youzhen

; APPLICANT: Xu, Yongyao

; APPLICANT: Hoersch, Sebastian

; APPLICANT: Monahan, John

; APPLICANT: Meyers, Rachel E.

; APPLICANT: Bast Jr., Robert C.

; APPLICANT: Hortobagyi, Gabriel N.

; APPLICANT: Pusztai, Lajos

; APPLICANT: Meric, Funda

; APPLICANT: Sahin, Aysegul

; APPLICANT: Mills, Gordon B.

; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT

QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 Db 1411 GACCTCAGGCTCTTCAGAACTGCAAGTAATCCGGGAGCAATTCGCACAATGCGGCC 1470  
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 Db 1471 TACTCGCTCAGCTCGCAAGGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGAA 1530  
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480  
 Db 1531 CTGGGAGTGGACTGGCCCTCATCCACATAACACCCACTCTGCTTCTGTGCACACGGTG 1590  
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 Db 1591 CCTTGGGACCACTCTTTCGGAACCGCCACCAAGCTCTGCTCCACACTGCCAACCGCCA 1650  
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 Db 1651 GAGGACGAGTGTGTGGCGAGGGCTTGGCTGCGCACAGCTGTGCGCCGAGGCGCACTGC 1710  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 Db 1711 TGGGGTCCAGGGCCCACTGTCAGCTGTCAACTGCGCCAGTTCCTTCGGGGCCAGGAGTGC 1770  
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
 Db 1771 GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCAGGGAGTATGTGTAATGCCAGCACTGT 1830  
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 Db 1831 TTGCGCTGCCACCTCAGTGTGAGCCCAAGATGGCTGACCTGACCTGTTTGGACCGGAG 1890  
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
 Db 1891 GCTGACCACTGTGTGGCTGTGCCACTATAAGGACCCCTCCCTCTGCGTGGCCCGCTGC 1950  
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
 Db 1951 CCCAGGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010  
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640  
 Db 2011 GCGCATGCGCAGCTTGCCCATCAACTGCACCACTCTCTGTGTGACCTGGATGACAAG 2070  
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
 Db 2071 GGCTGCGCGCGCGAGCAGAGCCAGCCCTCTGACGTCC 2109

; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER  
 ; FILE REFERENCE: MRI-038  
 ; CURRENT APPLICATION NUMBER: US/10/177,293  
 ; CURRENT FILING DATE: 2002-06-21  
 ; PRIOR APPLICATION NUMBER: US 60/299,887  
 ; PRIOR FILING DATE: 2001-06-21  
 ; PRIOR APPLICATION NUMBER: US 60/301,572  
 ; PRIOR FILING DATE: 2001-06-27  
 ; PRIOR APPLICATION NUMBER: US 60/306,501  
 ; PRIOR FILING DATE: 2001-07-18  
 ; PRIOR APPLICATION NUMBER: US 60/325,002  
 ; PRIOR FILING DATE: 2001-09-25  
 ; PRIOR APPLICATION NUMBER: US 60/362,585  
 ; PRIOR FILING DATE: 2002-03-05  
 ; PRIOR APPLICATION NUMBER: US 60/xxx,xxx  
 ; PRIOR FILING DATE: 2002-05-14  
 ; NUMBER OF SEQ ID NOS: 506  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 125  
 ; LENGTH: 4530  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-177-293-125

Alignment Scores:  
 Pred. No.: 0 Length: 4530  
 Score: 3628.00 Matches: 653  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 14 Gaps: 0

SEQ3 (1-653) x US-10-177-293-125 (1-4530)

QY 1 MetGluLeuAlaLeuCysArgTyrGlyLeuLeuAlaLeuLeuProGlyVala 20  
 DB 151 ATGGAGCTGGGGGCTGTGGCGTGGGGGCTCTCTCGCCCTTGTCCCGGAGGCC 210  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 DB 211 GCGAGCACCCAAAGTGTGCACCGGCACACATGAAGCTGGCGCTCCCTGCCAGTCCCGAG 270  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
 DB 271 ACCACCTGGACATGCTCCGGCCACTCTTACCAGGGCTGCCAGGTGGTGCAGGGAAACCTG 330  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
 DB 331 GAACCTACCTACCTGCCACCAATGCCAGCCTGTCTTCTGACAGATATCCAGAGGTG 390  
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 DB 391 CAGGGCTACGTGCTCATCGCTCACAAACCAAGTAGGAGGCTCCCACTGCAGAGGCTGGG 450  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 DB 451 ATTTGGGAGGACCCAGCTCTTTGAGGACAACCTATGCCCTGGCCGCTGCTAGACAATGGA 510  
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 DB 511 GACCGCTGACATACCAACCCCTGTACAGGGGCTCCCAAGAGGCTGCGGGAGCTG 570  
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160  
 DB 571 CAGCTTGAAGCTCAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCGAG 630  
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180  
 DB 631 CTCTGCTACAGGACACGATTTTGTGAAGGACATCTTCCACAAGAACCAACAGCTGGCT 690  
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 DB 691 CTACACTGATAGACACCAACCCCTCTCGGGGCTGCCACCCCTGTTCTCCGATGTGTAAG 750

QY ~ 201 GlySerArgCysTyrGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 DB 751 GGCTCCCGCTGCTGGGGAGAGATTCTCAGGATTCTCAGAGCTGACGCGCACTGTCTGT 810  
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
 DB 811 GCCGGTGGCTGTGCCCGCTGCAAGGGGCACTGCCCACTGACTGCTGCTCATGAGCAGTGT 870  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 DB 871 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGGCTGCTCCTCACTTCAACCA 930  
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 DB 931 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTACACACAGACACAGTTTGGAG 990  
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 DB 991 TCCATGCCCAATCCCGAGGGCGGTATACATTCCGGCCGAGCTGTGCTGACTGCTGCTGCC 1050  
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 DB 1051 TACAACCTACCTTCTTACGGACGTGGGATCCCTGCACCCCTGCTGCCCTGCTGCCCTGCACACCA 1110  
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 DB 1111 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAAGTGCAGCAACCCCTGTCGCCGA 1170  
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 DB 1171 GTGTGCTATGCTTGGGCATGAGCAGCTTGCAGAGGTGAGGGCAGTTTACCAGTGCCCAAT 1230  
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 DB 1231 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGGAGCCTGGGCAATTTCTCCCGGAGAGC 1290  
 QY 381 PheAspGlyAppProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
 DB 1291 TTTGATGGGACCCAGCCTCCACACACTGCCCGCTGCCAGCCAGACAGCAGCTCCAAGTGT 1350  
 QY 401 GluThrLeuGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420  
 DB 1351 GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGCCGCGGAGAGCCTGCT 1410  
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 DB 1411 GACCTCAGCGTCTTCCAGAACCTGCAAGTAAATCCGGGACGAATTTGCAAAATGGCGCC 1470  
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 DB 1471 TACTCGCTGACCTGTCAAGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGGAA 1530  
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480  
 DB 1531 CTGGCAGTGGACTGGCCCTCATCCACATAACACCCACCTCTGCTGCTGTCACACAGGTG 1590  
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 DB 1591 CCCTGGGACCACTCTTTCCGGAACCCGACCACTGCTGCTCCACACTGCCAAGCGGCCA 1650  
 QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 DB 1651 GAGGACAGGTGTGGGGCAGGGGCTGGCCCTGCCCTGCCAGCTGTCGCGCGGAGGCACTGC 1710  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 DB 1711 TGGGCTCCAGGGCCCACTGTCAGTGTCACTGACGAGCAGTTCTCTTCGCGGGCCAGGAGTGC 1770  
 QY 541 ValGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
 DB 1771 GTGGAGGAATGCCAGTACTGCAGGGGCTCCCGAGGAGTATGTGTAATGCCAGGCACTGT 1830

```
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db |||||
1831 TTGCCGTGCCACCTGAGTGTCAAGCCAGATGCTCAGTGACCTGTTTGGACCGGAG 1890
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys 600
Db |||||
1891 GCTGACCAAGTGTGTGGCTGTGCCCACTATAGGACCTCCCTTCTGCGTGGCCGCTGC 1950
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db |||||
1951 CCCAGGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db |||||
2011 GCGCATGCCAGCTTGCCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAG 2070
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db |||||
2071 GGCTGCCCGCGGAGCAGAGAGCCAGCCCTCTGACGTCC 2109
```

Search completed: October 17, 2003, 03:54:11  
Job time : 732.009 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 620.221 Seconds  
(without alignments)  
2842.104 Million cell updates/sec

Title: SEQ3

Perfect score: 3628

Sequence: 1 MELAALCRWGLLLALLPPGA.....CVDLDDKGPAPQAPASPLTS 653

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV-xlp

-MODEL-frame\_p2n.model -DEV-xlp  
-Q/cgn2\_l/USPTO.spool\_p/HOLLERAM480/runat\_15102003\_131912\_20526/app\_query.fasta\_1.4685  
-DB-N\_Geneseq\_19Jun03 -OFMT-fastap -SUFFIX-rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE-LOCAL -OUTFMT=ptc -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER-HOLLERAM480 -CGCN\_1\_1\_1758 -runat\_15102003\_131912\_20526 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3628	100.0	2385	18	AAT72725 Her2-GM-CSF immuno
2	3628	100.0	3600	21	AAA89736 Human HER-2/neu co
3	3628	100.0	3768	17	AAT40739 Human HER-2/neu Oncogene
4	3628	100.0	3768	20	AA019112 Human HER-2/neu on
5	3628	100.0	3768	21	AA094555 Human heregulin 2
6	3628	100.0	3768	22	AA023392 Human HER-2/neu pr
7	3628	100.0	3768	24	AB235744 Human ERBB2 polyu
8	3628	100.0	3768	24	ABX09877 Human ERBB2 CDNA
9	3628	100.0	3768	24	AA043935 Human HER2 antigen
10	3628	100.0	3768	24	AA043986 Human ERBB2 DNA SE
11	3628	100.0	3768	24	ABV78168 Human HER-2/neu pr
12	3628	100.0	3768	24	AA032743 Human HER-2/neu CD
13	3628	100.0	3768	24	ABA92250 Human HER-2/neu DN
14	3628	100.0	3768	24	ABK10730 Human polynucleoti
15	3628	100.0	3768	24	ABL91709 Human HER2 (Erbb2)
16	3628	100.0	3768	24	ABK14058 cDNA encoding the
17	3628	100.0	4472	21	AAA14812 Human tumour anti
18	3628	100.0	4473	19	ABQ76220 HER-2 nucleic acid
19	3628	100.0	4473	20	AAZ31071 Human gene express
20	3628	100.0	4473	24	ABZ34969 Human HER-2 DNA
21	3628	100.0	4473	24	AA038904 Human HER2 (Erbb2)
22	3628	100.0	4530	16	AA015853 Nucleotide sequenc
23	3628	100.0	4530	18	AA015853 Human HER2 gene
24	3628	100.0	4530	21	AAZ60815 Human-tyrosine kin
25	3628	100.0	4530	22	AA019731 Human gene express
26	3628	100.0	4530	24	ABZ35012 Breast carcinoma r
27	3628	100.0	4530	24	ABV94128 Human HER2-neu SEQ
28	3628	100.0	4530	24	ABN85585 Human CDNA differe
29	3628	100.0	4530	25	ABK83918 Breast cancer asso
30	3628	100.0	4530	25	ACC50139 Human Her2/Neu enc
31	3628	100.0	4530	25	ABQ83856 HER2 transgene pla
32	3628	100.0	9274	22	AAF24297 HER-2 transgene pl
33	3628	100.0	9274	24	AA043934 Human HER2 (Erbb2)
34	3628	100.0	9274	24	ABK14057 Sequence encoding
35	3585	98.8	4299	14	AAZ45086 DC8scfv-erbB2EC fu
36	3525	97.2	2871	21	AAZ50586 cDNA encoding huma
37	3458	95.3	3678	24	ABK86207 Extracellular port
38	3422	94.3	1872	11	AAQ06828 Rat neu promoter.
39	3106	85.6	3955	16	AA01590 Mouse Her-2/neu ex
40	3106	85.6	3955	21	AAA87553 Her-2/neu extracel
41	3104.5	85.6	2763	24	ABA92252 Mouse Her-2/neu CD
42	3104.5	85.6	2781	24	AAA87553 Nucleotide sequenc
43	3100.5	85.5	3771	22	AAH42210 Mouse Her-2/neu CD
44	3100.5	85.5	3771	22	ABA92251
45	3100.5	85.5	3771	24	ABA92251

# ALIGNMENTS

RESULT 1  
AAT72725  
ID AAT72725 standard; cDNA: 2385 BP.  
XX  
AC AAT72725;  
XX  
DT 17-SEP-1997 (first entry)  
XX  
DE Her2-GM-CSF Immunostimulant fusion protein DNA.  
XX  
KW Her2-GM-CSF; granulocyte macrophage colony stimulating factor;  
KW growth factor receptor; oncogene; immunostimulant; cancer;  
KW therapy; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers

```

FT CDS 11...2359
FT /*tag= a
FT /product= GM-CSF-Her2 fusion protein
FT mRNA 11...1969
FT /*tag= b
FT /product= Her2
FT mRNA 1970...1975
FT /*tag= c
FT /product= Leu-Glu linker
FT mRNA 1976...2359
FT /*tag= d
FT /product= GM-CSF
XX
XX WO9724438-A1.
XX
XX 10-JUL-1997.
XX
XX 23-DEC-1996; 96WO-US20241.
XX
XX 28-DEC-1995; 95US-0579823.
XX
XX (ACTI-) ACTIVATED CELL THERAPY INC.
XX
XX Laus R, Ruegg CL, Wu H;
XX
XX WPI; 1997-363674/33.
XX P-PSDB; AAW19764.
XX
XX Potent APC that activates T-cells to give multivalent cellular
XX immune response - can also induce a cytotoxic T-cell response in a
XX vertebrate subject
XX
XX Disclosure; Fig 8; 45pp; English.
XX
XX A nucleic acid molecule (AAT72725) codes for a fusion protein
XX (AAW19764) comprising granulocyte-macrophage colony stimulating
XX factor (GM-CSF) and Her2, a growth factor receptor that is
XX over-expressed in breast and ovarian cancer cells. It was
XX prepd. by PCR amplification of Her2 cDNA from a breast cancer
XX cell line and fusion to GM-CSF cDNA. Fusion expression vectors can
XX be used to transfect mammalian and insect cells. The Her2-GM-CSF
XX fusion protein is used to generate anti-Her2 immunity. Tumour
XX cells are eliminated by cytotoxic T lymphocytes activated in vivo
XX or in vitro by exposure to antigen-presenting cells exposed to the
XX fusion protein.
XX
XX Sequence 2385 BP; 488 A; 780 C; 677 G; 440 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 8.69e-230 Length: 2385
XX Score: 3628.00 Matches: 653
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 18 Gaps: 0
XX
XX SBQ3 (1-653) x AAT72725 (1-2385)
XX
XX Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
XX Db 11 ATGGAGCTGGCGGCTGTGCGGCTGGGGCTCTCTCGCCCTCTTGGCCCCCGGAGCC 70
XX
XX Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
XX Db 71 GCGAGCACCACAAAGTGCACCGGCACACATGAGCTGGGCTCCCTGCCAGTCCCGAG 130
XX
XX Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrglnGlyCysGlnValValGlnGlyAsnLeu 60
XX Db 131 ACCACCTGGACATGCTCCGCCACCTCTACCAGGCTGCCAGGTGGTGCAGGGAACCTG 190
XX
XX Qy 61 GluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
XX Db 191 GAACCTACCTACCTGCCCCACCAATGCCAGGCTGTCTTCTTGCAGGATATCCAGGAGGTG 250

```

```

Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 251 CAGGGCTACGTGCTCATCGCTCACAACTGAGGAGGAGTCCCACTGCAGAGCTGGG 310
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 311 ATTGTGGAGGACCCAGCTCTTTGAGGACAACATATGCCCTGGCCGTGCTAGACAATGA 370
Qy 121 AspproLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 371 GACCCGCTGAACAATACCCCTGTCCAGGGGCTCCCAAGAGAGCTGCGGGAGCTG 430
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 431 CAGCTTCGAAGCTCAGAGAGATCTTGAAGAGGGGTCTTGATCCAGCGAACCCTCAG 490
Qy 161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 491 CTCTGCTACCAGGACACGATTTTGTGAAGGACATCTTCCACAAGAACAACACAGCTGGCT 550
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 551 CTCACACTGATAGACACCAACCGCTCTCGGGCTTGCACCCCTGTCTCCGATGTGTAAG 610
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 611 GGCTCCCGCTGCTGGGAGAGATCTCTGAGAGTTGTCAGAGCTGAGCGCACTGCTGT 670
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 671 GCGGGTGGCTGTGCCCTGCAAGGGGCCACTGCCCACTGACTGCTCCATGAGCAGCTGT 730
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 731 GCTGCCGGCTGCAGGGGCCCAAGCACTCTGACTGCTGCTGCTGCTCCACTTCAACAC 790
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 791 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTACACACAGACACAGCTTTGAG 850
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 851 TCCATGCCCAATCCGAGGGCCGATATACATTCGGGCCAGCTGTGTGACTGCTGCTGCC 910
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 911 TACAACCTACCTTTCTACGAGCTGGGATCTGCACCTCTGCTGCCCTCGCAACACAA 970
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 971 GAGGTGACACAGAGGATGGAACACACACGCGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1030
Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1031 GTGTGCTATGCTGTGGCCTGAGCACTTGCAGAGGTGAGGGCAGTTACAGTGCCTCAAT 1090
Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1091 ATCCAGAGGTGTGCTGGCTGCAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGAGAGC 1150
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1151 TTTGATGGGACCCAGCCCTCCACACTGCCCGCTCCAGCCAGCAGCAGCTCCAGTGT 1210
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1211 GAGACTTGAAGAGATCAGAGGTACCTATACATCTCAGCATGGCGGAGCAGCTGCCT 1270
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1271 GACCTACGCTGTCTCCAGAACCCTGCAAGTAAATCGGGGAGCAATCTTGCAACATGGCGCC 1330

```

```
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTIPLeuGlyLeuArgSerLeuArgGlu 460
Db 1331 TACTCGTACCCCTCAAGGGCTGGCATCAGCTGGCTGGGCTCACTGAGGAA 1390
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1391 CTGGGACAGTGGACTGGCTCATCCACCATAACACCCACCTCTGCTTGGTGCACAGGTG 1450
QY 481 ProTTPAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1451 CCTGGGACCACCTCTTTCGGAAACCGCACCAAGTCTGTCTCCACACTGCCAACCGGCA 1510
QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuLeuCysAlaArgGlyHisCys 520
Db 1511 GAGGACGAGTGTGGGCGAGGCTGGCTGCCACCACTGTGGCGCGAGGCACTGC 1570
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1571 TGGGGTCCAGGCCCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1630
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1631 GTGGAGGATGCGGAGTACTGAGGGCTCCCGAGGATATGTGAATGCCAGGACACTGT 1690
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1691 TTGCCGTGCCACCTGAGTGTGAGCCGCCAGATGCTCAGTCACCTGTTTGGACCGGAG 1750
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1751 GCTGACCACTGTGTGGCTGTGCCCACTATAAGGACCTCCCTTCTGCGCGCGCTGC 1810
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu 620
Db 1811 CCCAGGGTGTGAACCTGACCTCTCTACATGCCCACTGGAAGTTCCAGATGAGGAG 1870
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 1871 GCGCATGCGCAGCTTGGCCCATCAACTGCACCCACTCTGTGTGGACCTGGATGACAAG 1930
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1931 GGCTGCCCGCGGAGCAGAGACCGACCCCTCTGACGTCC 1969

RESULT 2
AAA89736
ID AAA89736 standard; DNA; 3600 BP.
XX
AC AAA89736;
XX
DT 12-JAN-2001 (first entry)
XX
DE Human HER-2/neu coding sequence.
XX
KW Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
KW colon cancer; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3600
FT FT /*tag= a
FT FT /product= "HER-2/neu protein"
XX
PN WO200044899-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-US02164.
XX
PR 29-JAN-1999; 99US-0117976.
XX
```

```
PA (CORI-) CORIXA CORP.
PA (SMIK ) SMITHKLINE BEECHAM.
XX Cheever MA, Gheysen D;
XX WPI; 2000-505976/45.
DR P-PSDB; AAB211198, AAB21208.
XX
PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and
PT prostate cancers -
XX
PS Disclosure: Fig 15; 128pp; English.
XX
CC The present sequence encodes the human HER-2/neu protein. HER-2/neu is
CC a member of the tyrosine kinase family of receptor-like glycoproteins
CC and shows homology to the epidermal growth factor receptor (EGFR). It
CC probably plays a part in cell growth and/or differentiation. The
CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising
CC a HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation
CC domain may be used to treat or prevent cancer by eliciting or
CC enhancing an immune response to the HER-2/neu protein. It may be used
CC to treat malignancies such as breast, ovarian, colon, lung and
CC prostate cancers, and may be used as an antigen to vaccinate against
CC these neoplasias.
XX
SQ Sequence 3600 BP; 723 A; 1108 C; 1075 G; 694 T; 0 other;

Alignment Scores:
Pred. No.: 1.43e-229 Length: 3600
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

SEQ3 (1-653) x AAA89736 (1-3600)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATFGAGCTGGCGGCTTGTGGCGCTGGGGGCTCTCTCGCCCTCTTGGCCCCCGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGACGCCCAAGTGTGCCCGCACACATGAGCTGCGCTCCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCCAGCTGGACATGCTCGCCACCTCTACCAAGGCTGCCAGGTGTGTCAGGAAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 181 GAATCAGCTACCTGCCCAATGCCAGCTGTCTTCTCTCAGGATATCCAGGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProGlnArgLeuArg 100
Db 241 CAGGCTAGTGTCTATCGCTCACAAACCAAGTGGAGGAGGAGTCCCTCCAGAGGTGCGG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTTGGCAGGCACCCAGCTCTTTGAGGACAACCTATGCCCTGGCGCTGTAGACAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCGCTGAACAATACCACCTCTCACAGGGGCTTCCCAAGGAGGCTTCGGGAGGTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTTCCAGCCCTCACAGATCTTTGAAGAGGAGGGTCTTGATCCAGCGGAACCCCGAG 480
QY 161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCGTCTACAGGACACGATTTTGTGGAAGGACATCTTCCCAAGGACATCTTCCAAACAACACAGCTG 540
```

QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 DB 541 CTCACACTGATAGACACCAACCGCTCTCGGSCCTGCCACCCCTGTTCTCCGATGTGTAG 600  
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 DB 601 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTGTTCAGAGCCTGACGCGCACTGTCTGT 660  
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
 DB 661 GCCGGTGGCTGTGCCCTGCAAGGGCCACTGCCCCACTGACTGCTGCCATGACAGCTGT 720  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 DB 721 GCTCCCGCTGCACGGGCCCAACGACTCTGACTGCTGCCCTGCCCTCCACTTCAACCCAC 780  
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 DB 781 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTTACACACAGACACAGCTTTGAG 840  
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 DB 841 TCCATGCCCAATCCGAGGGCCGTATACATTCGGCCGACGCTGTGTGACTGTGCTGTGCC 900  
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 DB 901 TACAACCTACCTTTCTACGGAGCTGGATCTGCACTGCTGCTGCCCTGCCACCAACAA 960  
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 DB 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGAAAGCCCTGTGCCGA 1020  
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 DB 1021 GTGTGCTATGCTTGGCATGAGGACACTTGGAGAGGTGAGGCGAGTTACAGTGCCCAAT 1080  
 QY 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 DB 1081 ATCCAGGAGTTGTGCTGCAAGAAGATCTTTGGGAGCTGGCATTTCTGCCGGAGAGC 1140  
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
 DB 1141 TTTGATGGGACCCAGCGCTCCAACTGCCCCGCTCCAGCGAGAGAGCTCCAGTGT 1200  
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420  
 DB 1201 GAGACTTGGAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGAGAGCTGCCT 1260  
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 DB 1261 GACCTCAGCGTCTCCAGAACCTGCAAGTAATCCGGGGAGCAATTCGCACAATGGCGCC 1320  
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 DB 1321 TACTCGCTGACCTTGCAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGAA 1380  
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480  
 DB 1381 CTGGGAGTGGACTGGCCCTCATCCACCATTAACACCCACTCTGCTTCGTGCACACCGGTG 1440  
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 DB 1441 CCCTGGGACCAAGCTCTTTCCGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500  
 QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 DB 1501 GAGGACGAGTGTGGGCGAGGGCTTGGCCCTGCCACAGCTGTGCGCCGAGGGGCACTGC 1560  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 DB 1561 TGGGGTCCAGGGCCCCAGTGTGTCACTGACGACAGTTCCTTCGGGGCCAGGAGTGC 1620

541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
 DB 1621 GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT 1680  
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 DB 1681 TTGCCGTGCCACCTGAGTGTGAGCCAGATGGCTCAGTGACCTGTTTGGACCGGAG 1740  
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
 DB 1741 GCTGACCACTGTGGGCTGTGCCACTATAGACCCCTCCCTCTGCTGCGTGGCCGCTGC 1800  
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
 DB 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG 1860  
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAsPlys 640  
 DB 1861 GGCAGCATGCCAGCCTTGCCCATCACTGACCCACTCTCTGTGTGGACCTGGATGACAAG 1920  
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
 DB 1921 GGTCGCCCGCCGAGCAGAGAGCCGCTCTGACGTCC 1959  
 RESULT 3  
 AAT40739  
 ID AAT40739 standard; cDNA; 3768 BP.  
 XX AAT40739;  
 AC AAT40739;  
 XX 01-JAN-1997 (first entry)  
 DT XX  
 XX HER-2/neu oncogene.  
 DE XX  
 KW HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;  
 KW breast cancer; ovary cancer; colon cancer; lung cancer;  
 KW prostate cancer; genetic immunisation; tumour; vaccine; vector;  
 KW ss.  
 XX XX  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 1..3765  
 FT /\*tag= b  
 FT /note= "nucleotides 2026-3765 (claim 1) code for  
 HER-2/neu intracellular domain"  
 XX  
 PN W09630514-A1.  
 XX  
 PD 03-OCT-1996.  
 XX  
 PP 28-MAR-1996; 96WO-USO1689.  
 XX  
 PR 31-MAR-1995; 95US-0414417.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 XX Cheever MA, Disis ML;  
 XX WPI; 1996-455361/45.  
 XX P-PSDB; AAW01111.  
 DR  
 XX DNA encoding HER-2-neu poly:peptide(s) - used for prevention or  
 treatment of malignancies with which the HER-2/neu oncogene is  
 associated  
 PT  
 XX Claim 1; Page 49-56; 71pp; English.  
 PS  
 XX Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or  
 c-erbB2) protein (AAW01111). The oncogene is overexpressed in various  
 cancers, including breast, ovarian, colon, lung and prostate, and  
 appears to induce malignancies through quantitative mechanisms that  
 result from increased or deregulated expression of an essentially



CC normal gene product. Nucleotides 2026-3765 of the cDNA sequence  
CC code for the intracellular domain (Lys676-Val1255) of the HER-2/neu  
CC protein, which is useful for immunisation against malignancy.  
CC Nucleic acids can be used to direct expression of the intracellular  
CC domain in transformed host cells, or are used, alone or in a viral  
CC vector, for genetic immunisation of an animal.

SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:		
Pred. No.:	1.51e-229	Length:
Score:	3628.00	Matches:
Percent Similarity:	100.00%	Conservative:
Best Local Similarity:	100.00%	Mismatches:
Query Match:	100.00%	Indels:
DB:	17	Gaps:
		3768
		653

SEQ3 (1-653) x AAT40739 (1-3768)

QY	1	MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla	20
DB	1	ATGGAGCTGGCGGCTTGTGCGGTGGGGGTCTCTCTCCCTTGTGCCCGCGAGCC	60
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetIysLeuArgLeuProAlaSerProGlu	40
DB	61	GGAGACCCCAAGTGTGACCGGCACAGACATGAAGCTGGCTGCTCCCTGCCAGTCCCGAG	120
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlyAsnLeu	60
DB	121	ACCCACCTGGACATCTCCGCCACCTCTACACAGGGCTGCCAGGTGGTGGCAGGAACCTTG	180
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
DB	181	GAATCACTACCTGCCCAACCAATGCCAGCCTGTCTTCTGCAGGATATCCAGGAGGTG	240
QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
DB	241	CAGGGCTAGTGTCTATCGCTCACAACCAAGTAGGAGGTCTCCCACTGCAGAGCTCGCG	300
QY	101	IleValArgGlyThrGlnLeuPheIleuAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
DB	301	ATTGTGGAGGCCACCACTCTTGTAGGACAACTATGCCCTGGCGGTGTAGACAATGGA	360
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
DB	361	GACCCGCTGAACAATACCACCCCTGTACAGGGGCCCTCCCAAGAGSCCTCGCGGAGCTG	420
QY	141	GlnLeuArgSerLeuThrGluIleLeuIysGlyValLeuIleGlnArgAsnProGln	160
DB	421	CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCCAG	480
QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
DB	481	CTCTGCTACGAGCACACGATTTTGTGAAGAGCAATCTCCACAGAACCAACAGCTGGCT	540
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
DB	541	CTCACACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTCTAG	600
QY	201	GlySerArgCysTrpGlyLysSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
DB	601	GGCTCCCGCTGTGGGGAGAGATTCTGAGGATTGTTCAGAGCCTGCAGCGCACTGCTCTGT	660
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
DB	661	GGCGTGGCTGTGGCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT	720
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
DB	721	GCTGCCGGCTGCACGGGGCCCCAAGCACTCTGACTGCTGCGCTGCCCTCCCACTTCAACCA	780
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280

Db	781	AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACAAACAGACAGACGTTTGGAG	840
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	841	TCCATGCCCAATCCCGAGGGCGGTATACATTCCGGCGCCAGCTGTGTGACTGCTGTGCC	900
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACATACCTTTCTACGGAGCTGGGATCTGCACCTCGTCTGCCCTCTGCACAACCAA	960
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACAGCAGAGGATGGAACACACGCGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA	1020
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGGTCTGGGCATGGAGCACTTGGAGAGGTGAGGCGAGTTACCACTGCCAAT	1080
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGGAGATTGTGTGCTCCAGAAATCTTTGGAGCCTGGCATTTCTGCCGGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe	400
Db	1141	TTTGATGGGAGCCAGGCCTCCAACTGCCCGCTCCAGCCAGCAGCAGCTCCCAAGTGT	1200
Qy	401	GluThrLeuGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTGGAAGAGATCACAGTTACTATACATCTCACGNTGGCGGACAGCTGCCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCGCTTCCAGAACCTGCAGTAATCCGGGGACGAATCTTCACAAATGGCGCC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCTTGCAAGGCGTGGGCATCAGCTGGCTGGGGCTGCCTTCACCTGAGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGGCAGTGGACTGGCCCTTCATCCACCATTAACCCACCTCTGCTTCGTGCACACGGT	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACAGCTCTTTGGAAACCCGCCACCAAGCTCTGTCTCCACAGTGCCCAACCGGCCA	1500
Qy	501	GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGACAGATGTGTGGCGAGGGCGCTGGCCTGCCACACAGCTGTGCCGCCGAGGGCACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGTCACAGGCCACCCAGTGTGTCACTGACGACCATTTCTTCGGGGCCACAGGAGTGC	1620
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGATGTGTAATGCCAGGCACCTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCGGTGCCACCTGTAGTGTCAAGCCCAAGTGGCTCAGTGCACCTGTTTGGACCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACAGTGTGGGCGCTGTGCCACTATAAGGACCCCTCCCTCTGCGTGGCCCCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCAGCGGTGTGAACCTTGACCTCTCTACATGCCCATCTGGGAAGTTTCCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGCATGCCAGCTTGGCCCATCACTGACCCCACTCTCTGTGTGACCTGGATGACAAG	1920

QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
 DB 1921 GGCTGCCCGCGGAGAGAGCCAGCCCTCTGACGTCC 1959

## RESULT 4

AAAX01912  
 ID AAX01912 standard; DNA; 3768 BP.

XX AAX01912;

DT 21-APR-1999 (first entry)

XX Human HER-2/neu oncogene DNA.

DE HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;  
 KW malignancy; treatment; tumour; sa.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH 1..3768  
 CDS /\*tag= a

FT /product= "HER-2/neu"

FT /note= "oncogene"

FT 2026..3765

FT /\*tag= b

FT /note= "region which elicits immune response"

XX misc\_feature

XX US5869445-A.

XX 09-FEB-1999.

XX 01-APR-1996; 96US-0625101.

XX 01-APR-1996; 96US-0625101.

XX 17-MAR-1993; 93US-0033644.

XX 12-AUG-1993; 93US-0106112.

XX 31-MAR-1995; 95US-0414417.

XX (UNIW ) UNIV WASHINGTON.

XX Cheever MA, Disis ML;

XX WPI; 1999-152835/13.

XX P-PSDB; AAW92406.

XX Use of HER-2/neu polypeptides - for eliciting an immune response to  
 an HER-2/neu associated malignancy, particularly for treating or  
 preventing tumours

XX Claim la; Column 23-32; 26pp; English.

XX This sequence encodes the human HER-2/neu oncogene protein. A fragment  
 of this protein is used in a method for eliciting or enhancing an immune  
 response to HER-2/neu protein. The polypeptide can stimulate T cells and  
 B cells to produce an immune response to the HER-2/neu protein. The  
 method can be used for immunisation against a malignancy in which the  
 HER-2/neu oncogene is associated and in the treatment of an existing  
 tumour, or to prevent tumour occurrence or reoccurrence.

XX Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

## Alignment Scores:

Pred. No.: 1..51e-229 Length: 3768  
 Score: 3628.00 Matches: 653  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0

SEQ3 (1-653) x AAX01912 (1-3768)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProProGlyAla 20  
 DB 1 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCCTCTCGCTCCCTCTTGTGCCCGGAGCC 60  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 DB 61 GCGAGACCAAGTGTGCACCGGCACAGACATGAAGCTGGGCTGCCCTGCAGTCCCGAG 120  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
 DB 121 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGGAACCTG 180  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
 DB 181 GAACCTACCTACCTGCCCAATGCCAGCTGTCTCTCTCCAGGATATCCAGGAGGTG 240  
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 DB 241 CAGGGCTACGTGCTCATCGCTCACAAAGTAGGAGGCTCCCACTGCAGAGGCTCGG 300  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 DB 301 ATTGTGCGAGGACCCAGCTCTTTGAGACAACATATGCCCTGGCGTGTCTAGACAATGA 360  
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 DB 361 GACCGCTGAACAATACCACTCCAGCCCTGTACAGGGGCTCCCAAGAGGCTCGCGGAGCTG 420  
 QY 141 GluLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160  
 DB 421 CAGCTTGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGAACCCTCAG 480  
 QY 161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180  
 DB 481 CTCTGTACCAAGGACAGATTTGTGGAAGGACATCTTCACCAAGAACACACAGCTGGCT 540  
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 DB 541 CTCACACTGATAGACACCAACCGCTCTCGGGGCTGCCACCCCTGTCTCGGATGTGAAG 600  
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 DB 601 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGCAGAGCCTGCAGCCACTGTCTGT 660  
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
 DB 661 GCCGTGGCTGTGCCCGCTCAAGGGGCCACTGCCCACTGCTGCTGCATGACCACTGT 720  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 DB 721 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCCCTGCCCTCCACTTCAACCAC 780  
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 DB 781 AGTGGCATCTGTGAGCTGCACCTCCAGCCCTGGTCACTTACCAACACAGACACGCTTTGAG 840  
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 DB 841 TCCATGCCCAATCCCGAGGGCCCGGTATACATTCGGGCCACGCTGTGTGACTGCTGCC 900  
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 DB 901 TACAACCTACCTTTCTAGCGAGCTGGGATCTCGACCCCTGCTGCCCTCCGCAACCA 960  
 QY 321 GluValThrAlaGluaspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 DB 961 GAGGTGCACAGCAGAGGATGGACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA 1020  
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 DB 1021 GTGTGTATGGTCTGGGCATGGAGCACTTCGAGAGGTGAGGCGAGTTTACAGTGCACAT 1080  
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380



Db	1381	CTGGCGAGTGGACTGGCGCCTCATCCACCATAACACCCACCTCTGCTCTGCTGTCACACCGGTG	1444
QY	481	ProTTPAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCACCGGCCA	1500
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACGAGTGTGTGGCGAGGCGCTGGCCTGCCACCAAGCTGTGCGCCCGAGGCACTGC	1560
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC	1620
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGAAATGCCGAGTACTGCAAGGCGCTCCCGAGGAGTATGTGAATGCCAGCACTGT	1680
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCTTGAGTGTCAAGCCCAAGTGGCTCAGTGACCTGTTTTGGACCGGAG	1740
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrTlyAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCAAGTGTGGCCTGTGCGCACTATAAGGACCTCCCTCTCTGGTGGCCCGCTGC	1800
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleThrLysPheProAspGluGlu	620
Db	1801	CCGAGCGGTGTGAACCTGACCTCTCTCATGCCATCGCAAGTTTCAGATGAGGAG	1860
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGCATGCCAGCCTTGCCCATCACTGCACCCACTCCTGTGTGGAGCTGGATGACAAG	1920
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCC	1959
RESULT 6			
AAH23392			
ID	AAH23392 standard; DNA; 3768 BP.		
XX	AAH23392;		
DT	25-SEP-2001 (first entry)		
DE	Human HER-2/neu protein encoding DNA.		
XX	Antigen-presenting cell; immunogenic; immune response; HER-2/neu;		
KW	oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.		
XX	Homo sapiens.		
XX	Key		
FT	Location/Qualifiers		
FT	1..3768		
FT	/*tag= a		
FT	/product= "HER-2/neu protein"		
XX	WO200153463-A2.		
PN	26-JUL-2001.		
XX	19-JAN-2001; 2001WO-US01850.		
PF	21-JAN-2000; 2000US-0177545.		
XX	(CORI-) CORIXA CORP.		
XX	Cheever MA, Hand-Zimmermann S;		
XX	WPI; 2001-476112/51.		
DR	P-PSDB; AAB85458.		
XX			

PT New antigen-presenting cells, useful as vaccines for eliciting or  
 PT enhancing an immune response to HER-2/neu protein, particularly useful  
 PT for treating or preventing cancer, e.g. breast cancer

XX Claim 1: Page 41-46; 49pp; English.

XX The invention provides an isolated antigen-presenting cell, which  
 CC expresses at least an immunogenic portion of a polypeptide that produces  
 CC an immune response to HER-2/neu protein. The antigen-presenting cells are  
 CC useful as vaccines for eliciting or enhancing an immune response to  
 CC HER-2/neu protein, particularly in treating or preventing malignancies in  
 CC which the HER-2/neu oncogene is associated. Specifically, these are  
 CC useful for treating or preventing cancer, e.g. breast cancer, ovarian,  
 CC colon, lung or prostate cancers. The present sequence represents a DNA  
 CC encoding the human HER-2/neu protein (also known as p185 or c-erbB2).

SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

#### Alignment Scores:

Pred. No.: 1.51e-229 Length: 3768  
 Score: 3628.00 Matches: 653  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0

SEQ3 (1-653) x AAH23392 (1-3768)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20  
 DB 1 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCTCTCGCCCTTGTGCCCCCGAGCC 60  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 DB 61 GCGAGCCCAAGTGTGACCGGCACAGACATGAAGCTGCGGCTCCCTGCGAGTCCCGAG 120  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
 DB 121 ACCCACTGGACATGCTCGGCACCTTACCAAGGCTGCCAGGTGGTGCAGGGAACCTG 180  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80  
 DB 181 GAACATCACTACTCTCCCAACCAATGCCAGCTGTCTTCTTCCAGAGATATCCAGAGGTG 240  
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 DB 241 CAGGCTAGTGCTCTATGCTCACAAACCAAGTACAGGAGTCCCACTGCAGAGGTGCGG 300  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 DB 301 ATTGTGCGAGGCACCCAGCTCTTGAGGACAACTATGCCCTGGCGGTGCTAGACAATGGA 360  
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 DB 361 GACCCGCTGAACAATACCAACCTCTACAGGGGCTCCCAAGAGGCTGCGGAGGTG 420  
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160  
 DB 421 CAGCTTCGAGAGCTCACAGATCTTGAAGAGGGGTCTTATCCAGCGAACCCTGCTGCT 480  
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180  
 DB 481 CTCTGCTACCAAGACACGATTTTGGAGGACATCTTCCACAAGAACACCAACCACTGGCT 540  
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 DB 541 CTCACATGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGCTTCCGATGTGAAG 600  
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 DB 601 GGCTCCCGCTGTGGGGAGAGATTCTGAGGATTGTGAGAGCTTACAGGCTTACGCGCACTCTGT 660  
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGluGlnCys 240

DB 661 GCCGCTGGCTGTGCCCGCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 720  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 DB 721 GCTGCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCGCTGCTCCACTTCAACCCAC 780  
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 DB 781 AGTGCACTCTGTGACTGCATGCGCCAGCCCTGGTGCACCTACACACAGACACAGTTGAG 840  
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 DB 841 TCCATGCCCAATCCCGAGGCGGTATACATTTCGCGCCAGCTGTGTGACTGCTGCTGCC 900  
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 DB 901 TACAATACCTTTCTACGGAGCTGGGATCTCTACACCTCTGCTGCCCTGTCACACCA 960  
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 DB 961 GAGGTGACAGCAGAGGATGGAAACAGCGGTGTGAGAGTGCAGAGCCCTGTGCCCGA 1020  
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 DB 1021 GTGTGCTATGTCTGGCATGGAGCACTTTCGAGAGGTGAGGCACTTACCAGTGCCTAAT 1080  
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 DB 1081 ATCCAGGAGTTTGTCTGGCTGCAAGAGATCTTTTGGAGCTGGCAATTTCTTGGCGAGAG 1140  
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
 DB 1141 TTTGATGGGACCCAGCCTCCAACTGCCCCCTCCAGCCAGCAGCAGCTCCAGTGTGT 1200  
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420  
 DB 1201 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGCGCGAGACGCTGCT 1260  
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 DB 1261 GACCTCAGCGCTTCCAGAACTGCAAGTAATTCGGGGACGAAATTTCTGCACAAATGCGCC 1320  
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 DB 1321 TACTGCTGACCTGCAAGGCTGGCATCAGCTGCTGGGCTGCGCTCCTCAGGAGNA 1380  
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480  
 DB 1381 CTGGCAGTGGACTGGCCCTCATCCACATAACACCCACCTCTGCTTGTGTCACAGGTG 1440  
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 DB 1441 CCCTGGGACCACTCTTTTCGGAACCCGACCAAGCTGCTCCACACTGCCAACCGGCA 1500  
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 DB 1501 GAGGACGAGTGTGTGGGAGGGCTGGCTGCCACCACTGCTGCGCCGAGGAGGACATGC 1560  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 DB 1561 TGGGTGCCAGGCCCACTGCTGTGTAACATGACCCAGTTCCTTCTGGGGCCAGGAGTGC 1620  
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
 DB 1621 GTGGAGGAATGCCGAGTACTGTCAGGGGTCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680  
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 DB 1681 TTGCGCTGCCACCTTGTGAGTGTGAGGAGTGTGAGGAGTGTGAGTGTGAGTGTGAGTGT 1740  
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600

Db 1741 GCTGACCAGTGTGGCGCTGTGCCCACTATTAAGACCCCTCCCTTCTGCGTGGCCCGCTGC 1800

Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu 620

Db 1801 CCCAGCGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860

Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640

Db 1861 GGCCGATGCCAGCCTTGCCCATCACTGCAACCCACTCTCTGTGGACCTGGATGACAAG 1920

Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653

Db 1921 GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCC 1959

RESULT 7

ABZ35744

ID ABZ35744 standard; DNA; 3768 BP.

XX

AC ABZ35744;

DT 07-FEB-2003 (first entry)

XX

XX Human ERBB2 polynucleotide SEQ ID NO 52.

XX

KW Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;

KW protozoacide; gene expression; antisense; tumour; infection; Plasmodium;

KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;

XX Hepatitis C virus; human papilloma virus; gene; ds.

OS Homo sapiens.

XX

XX DE10100588-A1.

PN

XX 18-JUL-2002.

PD

XX

XX 09-JAN-2001; 2001DE-1000588.

PF

XX

XX 09-JAN-2001; 2001DE-1000588.

PR

XX (RIBO-) RIBOPHARMA AG.

PA

XX

XX Kreutzer R, Limmer S, Rost S, Hadwiger P;

PI

XX WPI; 2002-683450/74.

DR

XX

XX Inhibiting expression of target genes, useful e.g. for treating tumors,

PT by introducing into cells two double-stranded RNAs that are

PT complementary to the target -

XX

XX Claim 13; Page 38-39; 100pp; German.

PS

XX

CC The invention relates to inhibiting expression of a target gene in a cell

CC by introducing at least two oligoribonucleotides (dsRNAi and II), both

CC with a double-stranded (ds) structure of at most 49 sequential nucleotide

CC pairs. At least part of one strand (S1, S2) of the ds structures in each

CC of dsRNAi and II are complementary to regions in the target gene. The

CC method uses antisense inhibition of gene expression using double stranded

CC RNA inhibition (RNAi). The method is particularly used to treat tumours

CC on infections, especially by Plasmodium or viruses/viroids (pathogenic on

CC humans, animals or plants). The method provides more effective inhibition

CC of expression than known methods using a single dsRNA, even at very low

CC concentrations. When dsRNA has at least one unpaired nucleotide at the

CC end, stability (and thus effective concentration in the cell) is

CC improved and efficiency can be increased further by pretreating the cells

CC with interferon. The present sequence is that of a target DNA of the

XX invention.

XX

SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:

Pred. No.: 1.51e-229 Length: 3768

Score: 3628.00 Matches: 653

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 24 Gaps: 0

SEQ3 (1-653) x ABZ35744 (1-3768)

Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20

Db 1 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCCTCTCGCCCTTGTGCCCGGAGGCC 60

Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40

Db 61 GCGAGACCCCAAGTGTGCACCGGCACACATGAAGCTGGGCTTCCCTGCGCAGTCCCGAG 120

Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnVal 60

Db 121 ACCACCTGGACATGCTCCGCCACCTCTTACCAGGGCTGCCAGGTGGTGGAGGAACCTG 180

Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80

Db 181 GACTCACCTACCTGCCCAATGCCAGCGCTGCTTCTCTGCAGGATATCCAGGAGGTG 240

Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100

Db 241 CAGGGCTACGTGCTCATCGCTCACACCAAGTAGGAGGTCCCACTGCAGAGCTGCGG 300

Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120

Db 301 ATTGTGCGAGCACCCAGCTCTTTGAGGACAATATGCCCTGGCGGTGCTAGACAAATGGA 360

Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140

Db 361 GACCCGCTGAACAATACCCCTCTCACAGGGGCTCCCAAGAGGAGGCTGCGGGAGCTG 420

Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160

Db 421 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCACGAGGACCCCGAG 480

Qy 161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180

Db 481 CTCTGCTACCAGGACAGCATTTGTGGAAGGACATCTTCCACAGAAACACCACTGGCT 540

Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200

Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTTGCACCCCTGTTCTTCCGATGTGAAG 600

Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220

Db 601 GGCTCCCGCTGCTGGGGAGAGATTCTGAGGATTGTTCAGAGCCTGACGCGCCTGTCTGT 660

Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240

Db 661 GCGGTGGCTGTGCGCGCTGCAGAGGGCCACTGCCCACTACTGCTGCTGATGAGCACTGT 720

Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260

Db 721 GCTGCGGCTGCACGGGCCCAAGCACTCTGACTGCGCTGCGCTGCCCTCAACCAAC 780

Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280

Db 781 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTACACACAGACAGACGTTTGGAG 840

Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300

Db 841 TCCATGCCCAATCCGAGGGCCGGTATACATTGCGGCCACGCTGTGTGCTGCTGCTCCC 900

Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320

Db 901 TACAACCTACCTTTCTACGGACGTGGGATCTCGACCCCTCTGCTGCCCGCTGCACAACCA 960

Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340

Db 961 GAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCTGTGCCCGGA 1020

```
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1021 GTGTCTATGTCGTGGCATGGAGCATTGGAGGTGAGGAGGTACAGTGCCTAAT 1080
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1081 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTGGAGCCTGGCATTTCTGCCGAGAGC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
Db 1141 TTTGATGGGACCCAGCCTCCAACTGCTCCCGCTCCAGCCAGAGCAGCTCCAAGTGT 1200
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420
Db 1201 GAGACTCGGAAGAGATCACAGTTTACCTATCATCTCAGCATGCGCGAGAGCTGGCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1261 GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGAGCAATTCGCAATGGCGCC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1321 TACTCGCTGACCCCTGCAAGGCTGGGCATCAGCTGGCTGGGGCTCGCGTCACTGAGGAA 1380
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 CTGGCAGTGCACCTGCGCTCATCCACCATTAACACCCACTCTGCTTCGTGCACACGGT 1440
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCCGGAACCAAGCTCTTCCGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500
QY 501 GluAspGluCysValGlyLeuGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGCAGCAGTGTGGGGAGGGCTGGCCCTGCCACCAAGCTGTGGCCCGGAGGCACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGTCAGGCGCCACCCAGCTGTCTCACTGCAGCCAGCTTCCTTCGGGGCCAGGAGTGC 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 GTGGAGGAATGCCAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGGCGTGCCACCCCTGAGTGTGACGCCAGAAATGGCTCAGTGACCTGTGTTGGACCGGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys 600
Db 1741 GCTGACCACTGTGTGGCTGTGCCCACTATTAAGGACCTCTCCCTTCTGGCTGGCCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGGTGAAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1861 GGCGCATGCCAGCTTGGCCCATCACTGCACTGCACTGCACTGCTGTGTGGAGCTGGATGACAAG 1920
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCCGAGAGAGAGCCGCTCTGACGTCC 1959
RESULT 8
ABX09987
ID ABX09987 standard; DNA; 3768 BP.
XX
AC ABX09987;
XX
DT 23-JAN-2003 (first entry)
XX
```

Human ERBB2 DNA fragment SEQ ID 52.

Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental; prion; inhibition; human; ds.

Homo sapiens.

DE10100587-C1.

21-NOV-2002.

09-JAN-2001; 2001DE-1000587.

09-JAN-2001; 2001DE-1000587.

(RIBO-) RIBOPHARMA AG.

Kreutzer R, Limmer S, Rost S, Hadwiger P;

WPI; 2002-742209/81.

Inhibiting expression of target genes, e.g. oncogenes, in cells, by introduction of complementary double-stranded oligoribonucleotide, after treating the cell with interferon -

Disclosure; Page 43-44; 98pp; German.

This invention describes a novel method for inhibiting expression of a target gene by introducing into the cell that contains the target gene at least one oligoribonucleotide (dsRNA) that has a double-stranded (ds) structure of not more than 49 consecutive nucleotides (nt), where at least a segment of one strand of the ds structure is complementary with the target gene and the cells are treated with interferon before introduction of dsRNA. The method is used to inhibit expression of target genes, particularly oncogenes, cytokine genes, Id (not defined) protein genes; developmental or prion genes, or genes expressed in pathogenic organisms (particularly plasmidia) or in viruses or viroids (pathogenic in humans, animals or plants). Treating the cells with interferon greatly increases the extent to which dsRNA can inhibit expression of the target genes, and the effect is even greater when dsRNA are modified to increase their stability. ABX09936-ABX10075 represent gene fragments used to illustrate the method of the invention.

Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:

Pred. No.: 1.51e-229 Length: 3768

Score: 3628.00 Matches: 653

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 24 Gaps: 0

SEQ3 (1-653) x ABX09987 (1-3768)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20

Db 1 ATGAGCTGGCGGCTTGTGCGCTGGGGCTCTCTCCCTGCGCTCTTCCCGGAGGCC 60

QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40

Db 61 GCGAGCACCAAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 120

QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60

Db 121 ACCACCTGGACATGCTCCGCACTCTTACCAGGGCTGCCAGGTGGTGCAGGAAACCTG 180

QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80

Db 181 GAACCTACCTACCTGCCCCACCAATGCCAGCTTCTTCCTGCGAGGATATCCAGAGGTG 240

QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100



Db 241 CAGGGCTACGTCTCATCTGCTCACAACCAAGTAGGAGGTCTCCACTGCAGAGCTGCGG 300  
 Qy IleValArgGlyThrGlnLeuPheGluAspAsnThrAlaLeuAlaValLeuAspAsnGly 120  
 Db 301 ATTTGTGCGAGCACCCAGCTCTTTTGAGGACAACTATGCTTGGCCGTGTAGACAATGGA 360  
 Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 Db 361 GACCCGTGAACATACCAACCCCTGTACAGGGGCTCCAGAGAGGCTGCGGAGCTG 420  
 Qy 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyValLeuLeuGlnArgAsnProGln 160  
 Db 421 CAGCTTGAGCCCTCAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCGAG 480  
 Qy 161 LeuCysThrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180  
 Db 481 CTCTGCTACCCAGGACAGATTTTGTGAAGAGACATCTTCCACAAGAACCACTGGCT 540  
 Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGAAG 600  
 Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 Db 601 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTAGAGCCCTGACGCGCACTGTCTGT 660  
 Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
 Db 661 GCCGGTGGCTGCTGCCCTGCAAGGGCCACTGCCCACTGACTGCTCCATGACAGTGT 720  
 Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 Db 721 GCTCGCGCTGCACGGGCCCAACGACACTCTGACTGCTGGCTGCCCTCCACTTCAACCCAC 780  
 Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrThrAsnThrAspThrPheGlu 280  
 Db 781 AGTGGCATCTGTAGCTGCACTGCCAGCCCTGTGTCACCTACAAACACAGACACGTTTGAG 840  
 Qy 281 SerMetProAsnProGluGlyArgThrPheGlyAlaSerCysValThrAlaCysPro 300  
 Db 841 TCCATGCCCATCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGTGCC 900  
 Qy 301 TyrAsnThrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 Db 901 TACAACACTACCTTTCTAGCGAGTGGATCTCTGCACCCCTGCTGCTGCCCTGCACAAACAA 960  
 Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 Db 961 GAGGTGACAGCAGAGGATGGAAACACACGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA 1020  
 Qy 341 ValCysThrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 Db 1021 GTGTGCTATGTCTGGCATGGACACTTCGGAGAGGTGAGGCGAGTTACCACTGCCAAT 1080  
 Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 Db 1081 ATCCAGGAGTTTGTGTGCTGCAAGAAGATCTTTGGAGCCCTGGCATTTCTGCCGGAGAGC 1140  
 Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
 Db 1141 TTTGATGGGACCCAGCCTCCAAACACTGCCCGCTCCAGCCAGAGCAGCTCCCAAGTGT 1200  
 Qy 401 GluThrLeuGluGluIleThrGlyThrLeuTyrIleSerAlaTrpProAspSerLeuPro 420  
 Db 1201 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGCGCGGACAGCTGCT 1260  
 Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 Db 1261 CACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGAGCAAAATCTGCACAAATGGCGCC 1320  
 Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 Db 1321 TACTCGCTGACCCCTGCAAGGGCTGGGCATCAGCTGGGTGGGGCTGCGCTCACTAGGGAA 1380

Qy 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480  
 Db 1381 CTGGGCAGTGAGCTGGCCCTCATCCACATACACCACCCTCTGCTCTGTCACACGCTG 1440  
 Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 Db 1441 CCCTGGGACCACTCTTTTCGGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500  
 Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 Db 1501 GAGCAGCAGTGTGTGGCGAGGGCTGCGCTGCCACCAAGCTGTGCGCCGAGGACACTGC 1560  
 Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 Db 1561 TGGGGTCCAGGGCCACCCAGTGTGTAACATGCAGCCAGTTCTCTCGGGGCCAGGAGTGC 1620  
 Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
 Db 1621 GTGGAGGAATGCCGAGTACTGCAGGGCTCCCGAGGAGTATGTGAATGCCAGCACTGT 1680  
 Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 Db 1681 TTGCGGTGCCACCTGAGTGTACGCCCAAGTGTGCTCAGTGACCTGTTTGGACCCGAG 1740  
 Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
 Db 1741 GCTGACCACTGTGTGGCTGTGCCCACTATTAAGACCCCTCCCTTCTGCGTGGCCGCTGC 1800  
 Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
 Db 1801 CCCAGCGTGTGAAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG 1860  
 Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640  
 Db 1861 GGCCCATGCCAGCCTTGCCCACTCACTGCACCCACTCTGTGTGGACCTGGATGACAAG 1920  
 Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
 Db 1921 GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCC 1959  
 RESULT 9  
 AAD43935  
 ID AAD43935 standard; cDNA; 3768 BP.  
 XX AC AAD43935;  
 XX 13-DEC-2002 (first entry)  
 XX Human HER-2 cDNA.  
 DE Transgenic animal; transgenic; mammary gland cell; HER2; tumour;  
 KW cancer; therapy; apoptosis; cytostatic; human; gene; ss.  
 XX Homo sapiens.  
 OS Key Location/Qualifiers  
 XX CDS 1..3768  
 XX /\*tag= a  
 XX /product= "Human HER2 protein"  
 XX US2002035736-A1.  
 XX 21-MAR-2002.  
 XX 16-MAR-2001; 2001US-0811115.  
 XX 16-MAR-2000; 2000US-189844P.  
 XX (ERIC/) ERICKSON S.  
 XX (KING/) KING K.  
 XX (SCHW/) SCHWALL R.



PI Erickson S, King K, Schwall R;  
 XX WPI; 2002-401155/43.  
 DR P-PSDB; AAE26349.  
 XX  
 PT New transgenic non-human mammal that produces detectable levels of a  
 PT native human HER2 protein in its mammary gland cells, useful as tumor  
 PT models for testing HER2-directed cancer therapies, and for identifying  
 PT anticancer agents -  
 XX  
 PS Example 2; Page 24-26; 83pp; English.  
 XX  
 CC The invention relates to a transgenic non-human mammal that produces in  
 CC its mammary gland cells detectable levels of a native human HER2 protein  
 CC or its fragment. The transgenic animals are useful as tumour models for  
 CC testing HER2-directed cancer therapies, and for identifying anticancer  
 CC agents. The animals may also be used as source of cells which can be  
 CC immortalised in culture, in screening for compounds that have potential  
 CC as prophylactic or therapeutic treatments of diseases or disorders  
 CC involving expression of HER2. The anti-cancer molecules are useful for  
 CC inducing apoptosis or cell death of cancer cells. The present sequence  
 CC is human HER-2 cDNA.  
 XX  
 SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.51e-229 Length: 3768  
 Score: 3628.00 Matches: 653  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

SEQ3 (1-653) x AAD43935 (1-3768)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20  
 DB 1 ATGGAGCTGGGGCGCTTGGCGCTGGGGGCTCTCTCCCGCCCTTTCGCCCGGAGGCC 60  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 DB 61 GCGAGCACCACCAAGTGTGACCGGGCACACATGAGCTGGCGCTCCCTGCCAGTCCCGAG 120  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
 DB 121 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGAAACCTG 180  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
 DB 181 GAACCTCACTACCTGCCCAACCAATGCCAGCCTGCTCTCTCCAGGATATCCAGGAGGTG 240  
 QY 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 DB 241 CAGGGCTACGTGCTCATCGCTACAAACCAAGTGAAGGAGGCTCCCACTGCAGAGGTGCGG 300  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 DB 301 ATTGTGCCAGGACCCACCTCTTGGAGACAACTATGCTTGGCGTGTCTAGACAAATGGA 360  
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 DB 361 GACCCGCTGAACAATACCACCCCTGTACAGGGGCTCCCAAGGAGGCTCGGGAGGTG 420  
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160  
 DB 421 CAGCTTCGAAGCCCTACAGATCTTGAAGAGGGGTCTTTCATCCAGCGAAACCCCGAG 480  
 QY 161 LeuCysTyrGlnAspThrIleLeuTriPlysAspIlePheHisLysAsnAsnGlnLeuAla 180  
 DB 481 CTCGTCTACAGGACACGATTTTGTGGAGGACATCTTCCACAGAACCAACCACTGGCT 540  
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 DB 1621 GTGAGGAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680

DB 541 CTCACACTAGACACCAACCGCTCTCGGGCTCCACCCCTGTTCCTCGATGTGTAAAG 600  
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 DB 601 GGCTCCCCCTCTGGGGAGAGATTCTGAGGATTGTAGAGCTTACCGGCACTGTCTGT 660  
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
 DB 661 GCCGTGCTGTGCCCGCTGCAAGGGCCACTGCCACTGACTGTGCTGCATGAGCAGTGT 720  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 DB 721 GCTCGCGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGCTGCTCAACCAAC 780  
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 DB 781 AGTGGCATCTGTAGCTGCACCTGCCCGCCCTGGTCACTTCAACACACAGACAGTTGAG 840  
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 DB 841 TCCATGCCCAATCCCGAGGCGGTATACATTCGGCGCCAGCTGTGACTGCTGCTCC 900  
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 DB 901 TACAACCTACCTTTCTACGGAGCTGGGATCTCTGACCTCGTCTGCCCTGCACACCAA 960  
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 DB 961 GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAAGTGCAGCAAGCCCTGTGCC 1020  
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 DB 1021 GTGTGCTATGCTGGGCATGGAGCACTTGCAGAGGTGAGGGCAGTTACCACTGCCAAT 1080  
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 DB 1081 ATCCAGGAGTTTGTGCTGCAAGAAATCTTTGGAGCTGGCATTTCTGCCGGAGAGC 1140  
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
 DB 1141 TTTGATGGGACCCAGCCTCCAACTGCCCGCTCCAGCCAGACAGCACTCCAAGTGT 1200  
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420  
 DB 1201 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGCCCGGACAGCCTGCT 1260  
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 DB 1261 GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATCTGCACAATGGCGCC 1320  
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTyrLeuGlyLeuArgSerLeuArgGlu 460  
 DB 1321 TACTCGCTGACCCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTCGCTCTAGAGGAA 1380  
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480  
 DB 1381 CTGGCAGTGAGCTGGCCCTCATCCACCTAACACCCACCTCTGCTGCTGCACACGCTG 1440  
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 DB 1441 CCCTGGGACCAAGCTCTTTTCGGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500  
 QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 DB 1501 GAGGACGAGTGTGGGCGAGGGCTGGCTGCTGCCACCAAGCTGTGGCGCCGAGGCACTGC 1560  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 DB 1561 TGGGCTCCAGGGCCCAACCACTGTCACTGCAGCCAGTTCCTTCGGGGCCAGAGTGC 1620  
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
 DB 1621 GTGAGGAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680

QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 DB 1681 TTGCCGTGCCACCTGAGTGTACAGCCAGAAATGGCTCAGTGACCTGTTTGGACCGAG 1740  
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrIlysAspProPheCysValAlaAArgCys 600  
 DB 1741 GCTGACCAAGTGTGGGCTGTGGCCACTATAAGAGACCTCCCTTCTCGTGGCCGCTGC 1800  
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu 620  
 DB 1801 CCCAGCGGTGAAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGAG 1860  
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640  
 DB 1861 GCGCATGCGCAGCCTTGCCCATCAACTGCACCCACTCTCTGTGGACCTGGATGACAAG 1920  
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
 DB 1921 GGCTGCCCGCCGAGAGAGAGCCAGCCCTCTGACGTCC 1959

RESULT 10  
 AAD43986  
 ID AAD43986 standard; DNA; 3768 BP.  
 XX  
 AC AAD43986;  
 XX  
 DT 13-DEC-2002 (first entry)  
 XX  
 DE Human Her2 antigen DNA.  
 XX  
 KW Human; immune response; T-helper cell epitope; chitosan; CTL response;  
 KW vaccine; prostate cancer; breast cancer; Her2 antigen; cytostatic;  
 KW immunostimulant; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..3768  
 FT FT /\*tag= a  
 FT FT /product= "Human Her2 antigen"  
 FT sig\_peptide 1..69  
 FT FT /\*tag= b  
 FT mat\_peptide 70..2250  
 FT FT /\*tag= c  
 FT FT /product= "Mature human Her2 antigen"  
 XX  
 PN WO200234287-A2.  
 XX  
 XX  
 PD 02-MAY-2002.  
 XX  
 PF 26-OCT-2001; 2001WO-DK00705.  
 XX  
 PF 27-OCT-2000; 2000DK-0001606.  
 PR 03-NOV-2000; 2000US-245166P.  
 PR 18-JUN-2001; 2001DK-0000936.  
 XX  
 XX (PHAR-) PHARMEXA AS.  
 XX  
 XX Beier AM, Gautam A, Mouritsen S;  
 XX  
 DR WPI; 2002-463339/49.  
 DR P-PSDB; AAE26366.  
 XX  
 XX Inducing or enhancing an immune response against an antigen,  
 XX particularly cytotoxic T-lymphocyte responses, for treating or  
 XX ameliorating prostate or breast cancer, comprises administering the  
 XX antigen formulated with chitosan  
 XX  
 PS Claim 28; Page 85-90; 97pp; English.  
 XX  
 XX The invention relates to a method for inducing or enhancing an immune  
 CC response against a polypeptide antigen in an animal, including human.

CC The method comprises administering the polypeptide antigen or at least  
 CC one variant which includes at least one first T-helper cell epitope that  
 CC is foreign to the animal (foreign TH epitope) and is formulated with  
 CC chitosan. The polypeptide antigen is weakly immunogenic or non-  
 CC immunogenic. The invention is used as vaccine. The chitosan and  
 CC polypeptide antigen or its variant are useful in the preparation of an  
 CC immunogenic composition for inducing or enhancing an immune response,  
 CC particularly CTL response, against the polypeptide or protein antigen.  
 CC The method for inducing or enhancing an immune response is useful in  
 CC treating or ameliorating cancer, e.g. prostate or breast cancer. The  
 CC present sequence is human Her2 antigen DNA.

XX Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

# Alignment Scores:

Pred. No.: 1.51e-229 Length: 3768  
 Score: 3628.00 Matches: 653  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

SEQ3 (1-653) x AAD43986 (1-3768)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProProGlyAla 20  
 DB 1 ATGGAGCTGGCGGCTTGTGCCGCTGGGGCTCCTCTCGCCCTTTGCCCGGAGCC 60  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 DB 61 GCGAGCACCACCAAGTGTGCACCGCACAGACATGAAGCTGGGCTGCCCTGCCAGTCCCGAG 120  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnVal 60  
 DB 121 ACCACCTGGACATGCTCCGCCACTCTACAGGGCTGCCAGGTGGTGCAGGGAACCTG 180  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
 DB 181 GAACACCTACCTGCCCGCCACCAATGCCAGCTGTCTCTTCTGCAGGATATCCAGAGGTG 240  
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 DB 241 CAGGGCTACGTGTCTCATCGCTCACACCAAGTAGGAGGTGCCACATGCCAGAGCTCGCG 300  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 DB 301 ATGTGGGAGCACCACCTCTTTGAGGACAACATATGCCCTGGCCGTGTAGACAATGGA 360  
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 DB 361 GACCCGCTGAACAATACCAACCCCTGTACAGGGGCTCCCGCAGGAGGCTGCGGGAGCTG 420  
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160  
 DB 421 CAGCTTGAAGCCCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGSAACCCCGAG 480  
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180  
 DB 481 CTCTGTACCAGGACACGATTTGTGAAGAGCATCTTCCCAAGACACACACCTGGCT 540  
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 DB 541 CTCACACTGATAGACACCAACCCGCTCTCGGGCCCTGCCACCCCTGTCTTCCGATGTGAAG 600  
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 DB 601 GGCTCCCGCTGTCTGGGAGAGAGTTCYTAGGATTGTTCAGAGCTTCAGCGCACTGTCTGT 660  
 QY 221 AlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240  
 DB 661 GCCGGTGGCTGTGCCCGCTGCAGGGGCCACTGCCCACTGACTCTGCCATGACAGTGT 720  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260

```
Db 721 |GTCGGCGGTGACGGGCCCCAAGCACCTCTGCTGCGCTGCTCCACTTCACCCAC 780
QY SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 |AGTGGATCTGTAGCTGCATGCGCAGCCCTGGTCACTACACACAGACAGCTTTGAG 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 |TCCATGCCCAATCCCGAGGGCCGGTATACATTTCGGCGCCAGCTGTGTGACTGCCGTGCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 |TACAACCTACTTCTACGGACCTGGGATCTCTCACCTCTGCTGCCCTCGCACACACCAA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 |GAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAAGTGAGAGAGCCCTGTGCCCGA 1020
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValAlaValThrSerAlaAsn 360
Db 1021 |CTGTGCTATGCTGTGGCATGGACACTTGCAGAGAGTTCAGGCGAGTTACAGTGCCAAT 1080
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1081 |ATCCAGGAGTTTGTGCTGTGCAAGAGATCTTTGGGAGCTTGGCATTTCTGCCGGAGAGC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe 400
Db 1141 |TTTGATGGGACCCAGCTTCAACACTGCCCGCCCTCCAGCCAGAGAGCTTCAAGTGT 1200
QY 401 GluThrLeuGluGluIlePheGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1201 |GAGACTCTGGAAGATCATCAGGCTTACCTATACATCTCAGCATGCCGGACAGCTGCCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1261 |GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTTCTGCACAAATGGCGC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTyrPleuGlyLeuArgSerLeuArgGlu 460
Db 1321 |TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCAGTACAGGNA 1380
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 |CTGGCAGTGGACTGGCCCTATCCACCATACACCCACTCTGCTTCTGTGCACACGGTG 1440
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 |CCCTGGGACCACTCTTCGGAAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 |GAGGACGAGTGTGGCGGAGGGCTGGCTGCCACCACTGTGGCGCCGAGGCGACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 |TGGGTTCAGGGCCCAACCACTGTGTCAACTGCAGCCAGTTCTTCTCGGGCCAGGAGTGC 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 |GTGGAGGAATGCCGAGTACTTCAGGGGCTCCCGAGGAGATGTGAATGCCAGGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 |TTGCGCTGCCACCTGTAGTGTGAGCCCGACAGATGCTCAGTGACCTGTTTGGACCGGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys 600
Db 1741 |GCTGACCACTGTGTGGCTGTGCCCACTATAAGGACCTTCCCTTCTGCGTGGCCCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGlu 620
|
```

```
Db 1801 CCCAGCGGTGTGAACCTTGACCTCTCTATACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
|
Db 1861 GCGCATGCGCAGCTTGGCCCCATCAACTGCACCACTCTCTGTGTGGACCTGGATGACAAG 1920
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
|
Db 1921 GGCTGCCCGCCGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1959
RESULT 11
ABV78168
ID ABV78168 standard; DNA; 3768 BP.
XX
AC ABV78168;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human ERBB2 DNA SEQ ID NO 52.
XX
KW RNA inhibition; dsRNA1; gene expression inhibitor; oncogene; cytostatic;
KW virucide; protozoacide; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200255693-A2.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2002; 2002WO-EP00152.
XX
PR 09-JAN-2001; 2001DE-1000586.
PR 26-OCT-2001; 2001DE-1055280.
PR 29-NOV-2001; 2001DE-1058411.
PR 07-DEC-2001; 2001DE-1060151.
XX
PA (RIBO-) RIBOPHARMA AG.
XX
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
WPI; 2002-590671/63.
XX
PT Inhibiting expression of target gene, useful e.g. for inhibiting
PT oncogenes, by administering double-stranded RNA complementary to the
PT target and having an overhang
XX
PS Claim 10; Page 142-143; 203pp; German.
XX
CC The invention relates to inhibiting expression of a target gene (I) in a
CC cell by introducing an inhibitory RNA (dsRNA1) having a double-stranded
CC structure of at most 49 consecutive bases. At least part of one strand
CC (as1) of dsRNA1 is complementary to (I) and at least one end of dsRNA1
CC has an overhang of 1-4 nucleotides. The method is used to inhibit the
CC expression of a wide range of genes, e.g. oncogenes, cytokine genes etc.
CC in humans, also genes in Plasmodium or in viruses or viroids that are
CC pathogenic for humans, animals or plants. Introducing an overhang into
CC dsRNA greatly increases effectiveness for inhibiting gene expression,
CC both in vivo and in vitro and also increases stability and thus the
CC effective concentration inside the cell. The present sequence is that of
CC a gene related to the invention.
XX
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 1 51e-229 Length: 3768
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0
```

SEQ3 (1-653) x ABV78168 (1-3768)

Qy	1	MetGluLeuAlaLaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla	20
Db	1	ATGAGCTGGCGGCTTGTGCCTGGGGGTCCTCCTCGCCCTCTTTGCCCGGAGACC	60
Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	61	GCGAGCACCAAGTGTCACCGGCACAGCATGAAGCTGGCGCTCCCTGCAGTCCCAG	120
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrrGlnGlyCysGlnValValGlnGlyAsnLeu	60
Db	121	ACCACCTGGACATGCTCCGCCACTCTACAGGGCTGCCAGGTGGTCAGGAAACCTG	180
Qy	61	GluLeuThrTyrrLeuProThraSnaIaSerLeuSerPheLeuGlnAspIleGlnVal	80
Db	181	GAATCACCCTACCTGCCACCAATGCCAGCTGTCTTCTCCTGTCAGGATATCCAGAGGTG	240
Qy	81	GlnGlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	241	CAGGGCTACGTGTCATCGCTCAACAACAAGTAGAGCAGGTGCCACTGTCAGAGGCTCGG	300
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAspAsnGly	120
Db	301	ATTGTGGAGGCCACCCAGCTCTTTGAGACAACTATGCCCTGGCGGTGTAGACAAATGGA	360
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Db	361	GACCGCTGAACAATAACACCCCTGTACAGGGSCCTCCCCAGGAGGCTTCCGGGAGCTG	420
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160
Db	421	CAGCTTCGAAGCCTTCACAGAGATCTTGAAGAGGGGTCTTGATPCACAGCGAACCCCGAC	480
Qy	161	LeuCystyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
Db	481	CTCTGCTACAGGACACAGATTTGTGAAGAGCATCTTCCACAAGAACAACCACTGGCT	540
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	541	CTCACACTGATAGACACAACCGCTCTCGGSCCTGCCACCCCTGTTCTCCGATGTGTAAAG	600
Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	601	GGCTCCCGCTGTGGGAGAGATTCTGAGGATTGTTCAGAGCCTTCAGCGCACTGTCTGT	660
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	661	GCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGACAGTGT	720
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	721	GCTGCCGCTGCACGGGCCCCAAGCACTCTGACTGCTGCTGGCTGCCCTCCAACCAAC	780
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThrPheGlu	280
Db	781	AGTGGCATCTGTGAGCTTGCACTGCCAGCCCTGTGCACCTACCAACACAGACACGTTTGAG	840
Qy	281	SerMetProAsnProGluGlyArgTyrrPheGlyAlaSerCysValThrAlaCysPro	300
Db	841	TCCATGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGTCCC	900
Qy	301	TyrrAsnTyrrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACCTACCTTTCTACGGAGCTGGGATCTGCACCCCTGCTGTCGCCCTGCACAACCAA	960
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGlnLysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACAGCAGAGGTGAACAACACGCGTGTGAGAAGTGCACAGAACGCTGTGCCCGA	1020
Qy	341	ValCystyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGGTCTGGGCATGGAGCACTTCGGAGAGGTGAGGGCAGTTACACAGTGCCAAT	1080
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380

Db	1081	ATTCAGAGAGTTGCTGGCTGCACAGAGATCTTTGGAGCCTGTCATTTCTGCCGGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1141	TTTGATGGGACCAGCCTCCAACTGCCCCCTCCAGCCAGACAGCTCCAAGTGT	1200
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProaspSerLeuPro	420
Db	1201	GAGACTCTGAAGAGATCACAGGTACCTATACATCTCAGCATGGCCGACAGCCTGCCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgIleValGluLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCGCTTCCAGAACCTCGAAGTAATCCGGGGAGCAATCTGCACAATGGCGC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCTCGCAAGGCTGGGCATCAGCTGGCTGGGGCTGGCTCTACTGAGGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGCAGTGGACTGGCCCTCATCCACATAACACCCACCTCTGCTCTGTCACACGGTG	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGACCAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCTCCAACCCGGCCA	1500
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACAGTGTGTGGCGAGGGCTGGCCCTGCCACCACTGTGGCCCCGAGGGCACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGCCACCCAGTGTCACTGCAGGCAGTCTCTCGGGGCCAGGAGTGC	1620
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGAATGCCAGTACTGCGAGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCTGAGTGCAGCCCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys	600
Db	1741	GCTGACCAGTGTGTGGCTGTGGCCACTATAAGAGCCCTCCCTCTCGCTGGGCCCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProaspGluGlu	620
Db	1801	CCACGCGGTGTGAACCTGCACCTCTCTCATGCCATCTGGAAGTTCACGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGCATGCCAGCTTGCCCCCATCACTGCACCCACTCCTGTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGGTGCCCGCCCGAGAGAGAGCGAGCCCTCTCAGCGTCC	1959

RESULT 12	
AAD32743	
ID AAD3	
XX	
AC AAD3	
XX	
DT 01-J	
XX	
DE Huma	
XX	
KW Huma	
KW huma	
XX	
OS Homo	



```

Db 1201 GAGACTCTGGAGAGATCACAGGTTTACCTATACATCTCAGCATGCCGAGCAGCGCTGCCT 1260
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1261 GACCTCAGCGTCTCCAGAACCTGCAAGTATTCGGGGACGAATCTGCACAAATGGCGCC 1320
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1321 TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTCGCTCACTGAGGAA 1380
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 CTGGCAGTGGAGCTGGCCCTCATCCACATACACCCACCTCTGCTGTCACACGGTG 1440
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCTGGGACCAAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCA 1500
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGGACGAGTGTGTGGCGAGGCTGCGCTGCCACACAGCTGCGCCCGAGGACACTGC 1560
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGTCCAGGGCCCCACCACTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1620
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 GTGGAGGAATGCCGAGTACTCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCGTGCACCCCTGAGTGTGAGCCCGAGCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG 1740
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys 600
Db 1741 GCTGACCAAGTGTGGGCTGTGCCACATATAGGACCCCTCCCTTCTGCGTGGCCGCTGC 1800
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGTGTGAAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1861 GGCGCATGCCAGCCTTGCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAAG 1920
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCCCGGAGCAGAGAGCCAGCCCTCTGACGTCC 1959

RESULT 13
ID ABA92250
XX ABA92250 standard; cDNA; 3768 BP.
XX AC ABA92250;
XX AC ABA92250;
XX 17-JUN-2002 (first entry)
XX Human Her-2/neu cDNA.
XX Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
XX receptor; human; gene therapy; gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 1..3768
XX /*tag= a
XX /product= "Her-2/neu"
XX WO200212341-A2.
XX PN
XX PD 14-FEB-2002.

```

```

XX 03-AUG-2001; 2001WO-US24283.
XX 03-AUG-2000; 2000US-0632507.
XX (CORI-) CORIXA CORP.
XX PA (SMIK ) SMITHLINE BEECHAM BIOLOGICALS.
XX PI Cheever MA, Gheysen D;
XX WPI; 2002-241743/29.
XX DR P-PSDB; AAM51143.
XX Her-2/neu fusion protein for treating or preventing cancer by eliciting
XX or enhancing an immune response to the protein, has Her-2/neu
XX extracellular domain fused to Her-2/neu intracellular or
XX phosphorylation domain
XX Disclosure; Fig 15; 141pp; English.
XX The present sequence is that of human Her-2/neu oncogene, cDNA.
XX The cDNA encodes Her-2/neu (p185), an oncogenic self protein and
XX target for anti-cancer vaccines. The Her-2/neu gene is amplified
XX and p185 is overexpressed in a variety of cancers, including breast,
XX ovarian, colon, lung and prostate cancer. Her-2/neu is a member
XX of the tyrosine kinase family of receptor-like glycoproteins.
XX Its overexpression correlates with a poor prognosis in breast and
XX ovarian cancers. The invention provides Her-2/neu fusion
XX proteins, nucleic acids encoding them, viral vectors, and vaccines
XX comprising the fusion proteins or nucleic acid molecules. In
XX preferred fusion proteins, the extracellular domain of a Her-2/neu
XX protein is fused to a Her-2/neu intracellular domain or
XX phosphorylation domain (or its DeltapP fragment). An immune
XX response to Her-2/neu protein is elicited or enhanced by
XX administering the fusion protein in the form of a vaccine, or by
XX transfecting cells of an animal ex vivo with a nucleic acid
XX encoding the fusion protein, and delivering the transfected cells
XX to the animal. The fusion proteins, nucleic acids, and isolated
XX specific T-cells are useful for inhibiting the development of a
XX cancer, especially breast, ovarian, colon, lung or prostate cancer
XX in a patient. T cells that specifically react with a Her-2/neu
XX fusion protein can be used to remove tumour cells from a sample in
XX order to inhibit the development of cancer in a patient.
XX SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 1.51e-229 Length: 3768
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

SEQ3 (1-653) x ABA92250 (1-3768)
Qy 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGGAGCTGGCGGCGCTGTGCGCTGGGGGCTCTCTCGCCCTTGCCTCCCGGAGCC 60
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGCACCACCAAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCGAGTCCGAG 120
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValGlnGlyAsnLeu 60
Db 121 ACCCACCTGGACATGCTCCCGCACCTCTTACCAGGGCTGCCAGGTGGTCCAGGAAACCTG 180
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 181 GAATCCTACCTACCTGCCCAACCAATGCCAGCTGTCTCTCTCCAGGATATCCAGAGGTG 240
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100

```





PA (CORI-) CORIXA CORP.  
 XX Gaiger A, Cheever MA, Hand-zimmermann S;  
 PI WPI; 2002-280741/32.  
 DR P-PSDB; ANU77114.  
 XX Inhibiting haematological malignancy development by administering  
 PT polypeptide comprising immunogenic portion of Her-2/neu, polynucleotide  
 PT encoding the polypeptide, or antigen presenting cells expressing the  
 PT polypeptide  
 XX Disclosure; Page 66-71; 74pp; English.  
 XX The invention relates to a method for inhibiting development of  
 CC haematological malignancy in a patient by administering a polypeptide  
 CC comprising an immunogenic portion of Her-2/neu or a polynucleotide  
 CC encoding the polypeptide. Antigen presenting cells that express the  
 CC protein can also be administered. The sequences are used for inhibiting  
 CC development of haematological malignancy such as acute myelogenous  
 CC leukaemia (AML), chronic myelogenous leukaemia (CML), chronic lymphocytic  
 CC leukaemia (CLL), MDS, myelomas, Hodgkin's lymphoma and non-Hodgkin's  
 CC lymphoma. This sequence represents DNA encoding human Her-2/neu  
 CC polypeptide.  
 XX  
 SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1.51e-229 Length: 3768  
 Score: 3628.00 Matches: 653  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0  
 SEQ3 (1-653) x ABK10730 (1-3768)  
 QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProGlyAla 20  
 DB 1 ATGGAGCTGGCGGCTTGTGCGCGTGGGGCTCTCTCCCTCGCCCTTGTGCGCGGAGCC 60  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 DB 61 GCGAGACCCCAAGTGTGCACCGGCACACATGAGCTGGGCTCCCTGCGAGTCCCGAG 120  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
 DB 121 ACCCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGACAGGAACTG 180  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
 DB 181 GAACCTACCTACCTGGCCCAATGTCAGCGCTGTCTCTGCGAGGATATCCAGAGGTG 240  
 QY 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 DB 241 CAGGGCTACGCTCATCGCTCACACCAAGTAGGAGGCTCCACCTGCGAGGCTCGG 300  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 DB 301 ATTGTGGAGGCCACCGAGCTTTTGTAGGACAACATATGCCCTGGCCGTGCTAGACAATGA 360  
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 DB 361 GACCCGCTGAACAATACACCCCTGTACAGGGGCTCCCGAGGAGGCTCGCGGAGCTG 420  
 QY 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyValLeuIleGlnArgAsnProGln 160  
 DB 421 CAGCTTCAAGACCTCACAGAGATCTTGAAGAGGAGGGCTCTGATCCAGCGGAACCCCCAG 480  
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180  
 DB 481 CTCTGCTACCCAGGACGATTTTGTGGAAGGACATCTTCCACAGAACACACAGCTGGCT 540

QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 DB 541 CTCACACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAAG 600  
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 DB 601 GGCTCCCGCTGCTGGGAGAGAGTCTGAGGATTTGTGAGAGCCTGACGCGACTGCTGT 660  
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
 DB 661 GCCGGTGGCTGGCCGCTGCNAGGGCCACTGCCACTGACTGCTGCCATGAGCAGTGT 720  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 DB 721 GCTCCGGCTGCACGGGCCCAACGACTCTGACTGCTGCCCTGCCCTCCACTTCAACCCAC 780  
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 DB 781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGCTCACCTACACACAGACACGTTTGA 840  
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 DB 841 TCCATGCCCAATCCCGAGGGCCGTATACATTCGGCGCCAGCTGTGTGACTGCTGTC 900  
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 DB 901 TACAACCTACCTTTCTAGCGAGCTGGGATCTGTCACCTCGCTGCCCTGCACAAACCAA 960  
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 DB 961 GAGGTGACAGCAGAGGATGAAACACAGCGGTGTGAGAAGTGCACGAAGCCCTGTGCCGA 1020  
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 DB 1021 GTGTGCTATGCTGGCATGGGACACTTCCGAGAGGTGAGGGCAGTACCAGTGCACAT 1080  
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 DB 1081 ATCCAGGAGTTGCTGGCTGCAGAGAGATCTTTGGGAGCTTGGCATTTCTGCCGAGAGC 1140  
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnGlnValPhe 400  
 DB 1141 TTTGATGGGACCCAGCCTCCACACTGCCCCGCTCCAGCAGAGAGCTCCCAAGTGTGT 1200  
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpPrCAspSerLeuPro 420  
 DB 1201 GAGACTCTCGAAGAGATCAGAGTTACCTATACATCTCAGCATGCGCCGAGAGCTGCCT 1260  
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 DB 1261 GACCTCAGCGCTCTCCAGAACCTGCAAGTAAATCCGGGGAGGAATTCGCACATGGCGCC 1320  
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 DB 1321 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCCCTCACTCAGGGAA 1380  
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480  
 DB 1381 CTGGCAGTGGAGCTGGCCCTCATCCACCATTAACACCCACCTGCTTCTCGTGACACCGTG 1440  
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 DB 1441 CCCTGGGACCCAGCTCTTTCCGAACCCCGCACCAAGCTCTGTCTCCACACTGCCAACCGCCA 1500  
 QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 DB 1501 GAGGACAGTGTGTGGGCGAGGGCTGGGCTGCCACAGCTGTGCCGCCGAGGACATGC 1560  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 DB 1561 TGGGGTCCAGGGCCACCCAGTGTGTCACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1620  
 QY 541 ValGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560



Db 1621 CTGGAGGAATGCCGAGTACTGTCAGGGGCTCCCGCAGGAGTATGTGAATGCCAGGCACGTGT 1680  
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
Db 1681 TTGCGGTGCCACCCCTGAGTGTGAGCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG 1740  
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
Db 1741 GCTGACCACTGTGTGGCTGTGCCACTATAGGACCCCTCCCTTCTGGTGGCCGCTGC 1800  
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu 620  
Db 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860  
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAlaAspAspLys 640  
Db 1861 GCGCATGCCAGCCTTGGCCCATCACTGCAACCTCTCTGTGTGGACCTGGATGACAAG 1920  
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
Db 1921 GGCTGCCCGCGCGAGCAGAGCCAGCCCTCTGACGTCC 1959

RESULT 15  
ABL91709  
ID ABL91709 standard; DNA; 3768 BP.  
AC ABL91709;  
XX  
DT 28-MAY-2002 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 52.  
XX  
KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;  
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;  
KW cytostatic; virucide; protozoacide; antibacterial; ds.  
XX  
OS Homo sapiens.  
XX  
PN DE10100586-C1.  
XX  
PD 11-APR-2002.  
XX  
PF 09-JAN-2001; 2001DE-1000586.  
XX  
PR 09-JAN-2001; 2001DE-1000586.  
XX  
PA (RIBO-) RIBOPHARMA AG.  
XX  
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;  
XX  
XX WPI; 2002-270454/32.  
DR  
XX  
PT Inhibiting gene expression in cells, useful for e.g. treating tumors,  
PT by introducing double-stranded complementary oligoRNA having unpaired  
PT terminal bases  
XX  
PS Claim 13; Page 41-42; 104pp; German.  
XX  
CC The invention relates to a method for inhibiting expression of a target  
CC gene (ABL91658-ABL91797) in a cell by introducing at least one  
CC oligoribonucleotide that has a double-stranded structure consisting of at  
CC most 49 sequential nucleotide pairs, with at least part of one strand  
CC complementary with the target gene and has at least one end a  
CC single-stranded segment of 1-4 nt. The method provides  
CC oligoribonucleotides for antisense inhibition of gene expression useful  
CC e.g. for treating tumours but the oligoribonucleotides may also be  
CC directed against genes present in pathogens (e.g. Plasmodium or  
CC viruses/viroids, pathogenic on humans, animals or plants) or against  
CC cytokine, id, developmental or prion genes. The method provides more  
CC effective inhibition of gene expression than use of known  
CC oligonucleotides, probably because the unpaired overhang increases  
CC stability and thus intracellular concentration.

XX  
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;  
Alignment Scores:  
Pred. No.: 1.51e-229 Length: 3768  
Score: 3628.00 Matches: 653  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0  
SEQ3 (1-653) x ABL91709 (1-3768)  
QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20  
Db 1 ATGGAGCTGGGGCCCTTGTGCCCTGGGGGCTCTCTCGCCCTCTTGGCCCCCGGAGCC 60  
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
Db 61 GCGAGCACCCCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 120  
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu 60  
Db 121 ACCCACCTGGACATGCTGCCACCTCTACCAAGGCTGCCAGGTGGTGCAGGGAACCTG 180  
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
Db 181 GAATCACCCTACCTGCCACCACCAATGCCAGCCTGTCTCTCTGCAGATATCCAGGAGGTG 240  
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
Db 241 CAGGCTACGTGCTCATCGCTCACAAACCAAGTGAGCAGGTCCCCTGCAGAGGTGCGG 300  
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
Db 301 ATTGTGCGAGGCGCCAGCTCTTTGAGGACAACTATGCCCTGGCGCTGTAGACAATGA 360  
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
Db 361 GACCCGCTGAACAATACCACCCCTGTACAGGGGGCTCCCCAGGAGGCTTCCGGAGCTG 420  
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160  
Db 421 CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTTCATCCAGCGGAACCCCG 480  
QY 161 LeuCysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnAsnGlnLeuAla 180  
Db 481 CTCTGCTACCGGACACGATTTTGTGAAGGACATCTTCCACAAGAACAACCAAGCTGGCT 540  
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTTCTCCGATGTGAAG 600  
QY 201 GlySerArgCysTrpGlyLysSerGluAspCysGlnSerLeuThrArgThrValCys 220  
Db 601 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTCACAGAGCTTGCAGCGCACTCTCTGT 660  
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
Db 661 GCCGCTGCTGTGCCCTGCAAGGGGCACCTGCCCACTGACTGTGTCATGAGCAGTGT 720  
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
Db 721 GCTGCCGCTGCACGGGCCCCAAGCACTCTGACTGCTGGCTGCTCCACTTCAACCCAC 780  
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
Db 781 AGTGGCATCTGTGAGCTGCATGCCCGCCCTGGTCCACTACACACAGACACACGTTTTCAG 840  
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
Db 841 TCCATGCCCAATCCGAGGCGCGGTATACATTGGCGCGCAGCTGTGTGACTGCTGTGCC 900

Search completed: October 15, 2003, 23:49:41  
Job time : 669.221 secs

```
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 TACAACTACCTTTCTACGAGAGTGGGATCTCTGCACCTCGCTGCTGCCCTGCACAGCAA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCACAAAGCCCTGTGCCCGA 1020
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValAlaValThrSerAlaAsn 360
Db 1021 GTGTGCTATGCTTGGGCATGGACACTTGGAGAGGTGAGGGCAGTTACCAGTGCCAAT 1080
QY 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1081 ATCCAGGAGTTGTGGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGGAGAGC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1141 TTTGATGGGACCCAGCCTCCAACTGCCCCGCTCCAGCAGACAGCTCCAAGTGT 1200
QY 401 GluThrLeuGluGluIleThrGlyTyrIleSerAlaIleProAspSerLeuPro 420
Db 1201 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGGCCGGACAGCCTGCCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1261 GACCTCAGCGCTTCCAGAACCTCAAGTAATCCGGGAGCAATCTGCACAAATGGCGCC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1321 TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTGGCGCTCACTGAGGGAA 1380
QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 CTGGGCAGTGACCTGGCCCTCATCCACCAACACCCACCTCTGCTTCGTGCACACCGTG 1440
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCCTGGGACCAAGCTCTTTCGGAACCCGACCAAGCTCTGTCCACACTGGCCAAACCGCCA 1500
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGGACGAGTGTGGGCGAGGGCTGGCTGCCCTGCCAGCTGTGCCCGCCGAGGCACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGCTCCAGGGCCACCCAGTGTCAACTGCAGCCAGTTCTTCCGGGGCCAGGAGTGC 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCCGTGCCACCTGAGTGTACGCCCAAGTGGCTCAGTGACCTGTTTGGACCGGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrIlyAspProPheCysValAlaArgCys 600
Db 1741 GCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCCCTCCCTCTCGTGGCCCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 1861 GGCGCATGCCAGCCTTGCCCCCACTCACTGCCACCTCCCTGTGTGGACCTGGATGACAAG 1920
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCC 1959
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:09:13 ; Search time 9403.76 Seconds  
(without alignments)  
2840.777 Million cell updates/sec

Title: SEQ3  
Perfect score: 3628  
Sequence: 1 MELALCRWGLLALLPPGA.....CVLDLDDKGPAAEQRASPLTS 653

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -P/HOLLERAN480/runat\_15102003\_131912\_20536/app\_query.fasta\_1.4685  
-Q/cgn2\_1/USPTO\_spoel\_P/HOLLERAN480/runat\_15102003\_131912\_20536/app\_query.fasta\_1.4685  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 -CGN\_1\_1\_22724 -runat\_15102003\_131912\_20536 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3628	100.0	2385	6	AR082744	Sequence AR082744
2	3628	100.0	2385	6	AR099963	Sequence AR099963
3	3628	100.0	2385	6	AR143949	Sequence AR143949
4	3628	100.0	3768	6	AR034479	Sequence AR034479
5	3628	100.0	3768	6	AX060704	Sequence AX060704
6	3628	100.0	3768	6	AX201817	Sequence AX201817
7	3628	100.0	3768	6	AX380923	Sequence AX380923
8	3628	100.0	3768	6	AX384604	Sequence AX384604
9	3628	100.0	3768	6	AX465456	Sequence AX465456
10	3628	100.0	3768	6	AX467229	Sequence AX467229
11	3628	100.0	3768	6	AX481438	Sequence AX481438
12	3628	100.0	4473	6	AR080259	Sequence AR080259
13	3628	100.0	4473	6	AR167390	Sequence AR167390
14	3628	100.0	4473	9	HSEB2R	Sequence X03363 Human c-erb
15	3628	100.0	4530	6	AR202597	Sequence AR202597
16	3628	100.0	4530	6	AR283481	Sequence AR283481
17	3628	100.0	4530	6	AX282577	Sequence AX282577
18	3628	100.0	4530	6	AX587649	Sequence AX587649
19	3628	100.0	4530	6	AX644071	Sequence AX644071
20	3628	100.0	4530	6	BD005474	Sequence BD005474 Cellular
21	3628	100.0	4530	6	I21124	Sequence I21124
22	3628	100.0	4530	6	I59745	Sequence I59745
23	3628	100.0	4530	9	HUMHER2A	Sequence M11730 Human tyros
24	3628	100.0	9274	6	AX060703	Sequence AX060703
25	3525	97.2	2871	6	AX023363	Sequence AX023363
26	3471	95.7	1872	6	AR181000	Sequence AR181000
27	3458	95.3	3678	6	AX505114	Sequence AX505114
28	3324.5	91.6	3780	4	AB008451	Sequence AB008451 Canis fam
29	3110.5	85.7	4727	10	AY116182	Sequence AY116182 Rattus no
30	3106	85.6	3955	6	AX380924	Sequence AX380924
31	3106	85.6	3955	6	I21129	Sequence I21129
32	3106	85.6	3955	6	I59750	Sequence I59750
33	3106	85.6	3955	10	RNNEUR	Sequence X03362 Rat mRNA fo
34	3104.5	85.6	2763	6	AX380942	Sequence AX380942
35	3104.5	85.6	2781	6	AX380944	Sequence AX380944
36	3100.5	85.5	3771	6	AX189662	Sequence AX189662
37	3100.5	85.5	3771	6	AX380925	Sequence AX380925
38	3100.5	85.5	4694	10	BC046811	Sequence BC046811 Mus muscu
39	3100.5	85.5	4695	10	BC053078	Sequence BC053078 Mus muscu
40	3095	85.3	4062	10	HAMNEU	Sequence D16295 Mesocricetu
41	1874.5	51.7	1316	9	AF177761	Sequence AF177761 Homo sapi
42	1706.5	47.0	157216	9	AC087491	Sequence AC087491 Homo sapi
43	1706.5	47.0	161815	9	AC079199	Sequence AC079199 Homo sapi
44	1706.5	47.0	168585	9	AC040933	Sequence AC040933 Homo sapi
45	1706.5	47.0	198008	2	AC142197	Sequence AC142197 Homo sapi

ALIGNMENTS

RESULT 1



```

Db      1751  |||||||GCTACCAAGTGTGGCTGTGCCCACTATAAGGACCTCCCTTCTGGCTGGCCGCTGC 1810
Qy      601  ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db      1811  |||||||CCCAGCGGTGGAACCTGACCTCTCCACATGCCCATCTGGAAGTTCCAGATGAGGAG 1870
Qy      621  GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db      1871  |||||||GGCGCATGCCAGCTTGGCCCATCAACTGACCCACTCTGTGTGGACCTGGATGACAAG 1930
Qy      641  GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db      1931  |||||||GGCTGCCCGCGGAGCAGAGAGCCAGCCCTCTGACGTCC 1969

RESULT 2
LOCUS      AR099963      2385 bp      DNA      linear      PAT 14-FEB-2001
DEFINITION Sequence 3 from patent US 6080409.
ACCESSION  AR099963
VERSION     AR099963.1  GI:12810411
KEYWORDS   .
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 2385)
AUTHORS     Laus, R., Ruegg, C. Landon, and Wu, H.
TITLE       Immunostimulatory method
JOURNAL     Patent: US 6080409-A 3 27-JUN-2000;
FEATURES    Location/Qualifiers
             source          1..2385
                        /organism="unknown"
BASE COUNT  488 a 780 c 677 g 440 t
ORIGIN

```

```

Alignment Scores:
Pred. No.:      1.7e-226      Length:      2385
Score:           3628.00      Matches:      653
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:               6           Gaps:      0

```

SEQ3 (1-653) x AR099963 (1-2385)

```

Qy      1  MetGluLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db      11  ATGGAGCTGGCGGCTTGTGGCGTGGGGGTCTCTCTCGCCCTCTTGGCCCGGAGGCC 70
Qy      21  AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db      71  GCGAGCACCCCAAGTGTGCACCGGCACACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 130
Qy      41  ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValGlnGlyAsnLeu 60
Db      131  ACCCACTGGACATGCTCCGCGACCTCTACCAGGGGTGCCAGTGGTGGTGCAGGGAAACCTG 190
Qy      61  GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db      191  GAACATCACTACCTCCCAACCAATGCCAGCTGTCTCTTCTGAGGATATCCAGGAGGTG 250
Qy      81  GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      251  CAGGCGCTAGCTGCTCATCGCTCACAAACCAAGTAGGCGAGGTCCCACTGCAGAGGCTGCGG 310
Qy      101  IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      311  ATTTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCTCCCTGGCGTGTCTAGACAAATGA 370
Qy      121  AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db      371  GACCCGTGAACAATACCACCCCTGTACAGGGGCCCTCCCGAGGAGGCTCGGGAGCTG 430

```

```

Qy      141  GlnLeuArgSerLeuThrGluLeuLeuLysGlyValLeuLeuGlnArgAsnProGln 160
Db      431  CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGAGGGGTCTTGATCCAGCGGAACCCCGAC 490
Qy      161  LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
Db      491  CTCTGCTACCAAGGACACGATTTTGTGGAAAGGACATCTTCCACAAAGAACACAGCTGGCT 550
Qy      181  LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      551  CTCACATGATAGACACCAACCGCTCTCGGGCGCTCCACCCCTGTTCTCCGATGTGTAAAG 610
Qy      201  GlySerArgCysTrpGlyLysSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      611  GGCTCCCGCTGCTGGGGAGAGATTCTGAGGATTGTGAGAGCTCAGAGCTCAGCGCACTCTGT 670
Qy      221  AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db      671  GCCGGTGGCTGTGCCCGCTGCAAGGGCCACTGGCCCACTGACTGTCTGCCATGAGCAGTGT 730
Qy      241  AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db      731  GCTGCCGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGGCTGCTCCTCACTTCAACCCAC 790
Qy      261  SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db      791  AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTGCACCTACAACACAGACACAGTTTGA 850
Qy      281  SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db      851  TCCATGGCCCAATCCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGTGCC 910
Qy      301  TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db      911  TACAACCTACCTTCTACGGAGCTGGGATCTTCGACCCCTCGTCTGCTGCCCTGCACAAACCA 970
Qy      321  GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db      971  GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAAGTGAGCAAGCCCTGTGCCCGA 1030
Qy      341  ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db      1031  GTGTGCTATGCTTGGGCATGGAGCACTTGCAGAGGTGAGGGCAGTTACCAGTGCCTCAAT 1090
Qy      361  IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db      1091  ATCCAGGAGTTTGTGGCTGCAAGAGATCTTTGGAGCCCTGGCATTTCTGCCCGGAGAGC 1150
Qy      381  PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db      1151  TTTGATGGGACCCAGCCCTCCAACACTGCCCGCTCCAGCCAGACAGCAGCTCCAAGTGT 1210
Qy      401  GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db      1211  GAGACTCTCGAAGAGATCACAGGTACCTATACATCTCAGCATGTCGCCGAGCAGCTGGCT 1270
Qy      421  AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db      1271  GACCTCAGCTCTCCAGAACCTGCAAGTATATCCGGGAGCAGATTTCTGCACAAATGGCCGC 1330
Qy      441  TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db      1331  TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTCGCTCCTACTGAGGAA 1390
Qy      461  LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db      1391  CTGGCAGTGACTGGCCCTCATCACATAACACCCACCTCTGCTGCTGCTGCACACCGGTG 1450
Qy      481  ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db      1451  CCCTGGGACCAAGCTCTTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCCAACCGGCA 1510
Qy      501  GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520

```

Db	1511	GAGGACAGTGTGGCGGAGGCGCTGGCGTGCACACAGTGTGGCGCGGAGGCACTGC	1570
QY	521	TIPGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1571	TGGGGTCCAGGCGCCACCCAGTGTGCACTGCAGCCAGTTCCTTCGGGCGCCAGGAGTGC	1630
QY	541	ValGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1631	GTGAGGAATGCCAGTACTGCAGGGCTCCCCAGGAGTATGTGAATGCCAGGCACTGT	1690
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1691	TTCCCTGGCACCCCTGAGTGTACGCCCGCAGAAATGGCTCAGTACCTGTTTGGACCGGAG	1750
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1751	GCTGACCAAGTGTGGCGTGTGCCCACTATAGACCCCTCCCTTCGCGTGGCCCGCTGC	1810
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1811	CCCAGCGGTGAAACCTGACCTCTCTACATGCCCATCTGTGAAGTTTCCAGATGAGGAG	1870
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1871	GGCGCATGGCAGCCTTGGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	1930
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1931	GGCTGCCCGCCGAGCAGAGCAGCCCTCTGACGTCC	1969
RESULT 3			
AR143949			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			
SEQ3 (1-653) x AR143949 (1-2385)			
QY	1	MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla	20
Db	11	ATGGAGCTGGCGGCTTGTGGCGCTGGGGCTCTCTCTCGCCCTCTTGGCCCGCGGAGCC	70
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	71	GCAGACCCCAAGTGTGCACCGGCAGACATGAAGTGTGGCTCCCTGCGAGTCCCGAG	130
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60
Db	131	ACCCACTGGACATGTCCGCCACCTCTACAGGGCTGCGAGGTGTGGCGGGAACCTG	190
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	191	GAACTCACCTACCTGCGCCCAATGCCAGCTGTCTCTCTCTGAGGATATCCAGGAGGTG	250
QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	251	CAGGGTACGTGCTCATCTCACAACCAAGTGAGCAGGTCCCACTGCAGAGCGTCGG	310
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	311	ATTGTCCGAGGACCCAGCTCTTTGAGGAACTATGCCCTGGCCCTGTAGACAAATGGA	370
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Db	371	GACCCGCTGAACAATACCACCCCTGTACAGGGGCTCCCCAGGAGGCGCTGCGGAGCTG	430
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160
Db	431	CAGCTTCGAAGCCTCAGAGATCTTTGAAGGAGGGTCTTGATCCAGCGGAACCCCCAG	490
QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla	180
Db	491	CTCTGCTACAGGACACAGATTTGTGGAAGGACATCTTCACAAAGAAACACAGCTGGCT	550
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	551	CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTCTAAG	610
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	611	GGCTCCCGCTGCTGGGAGAGATTCTGAGGATTGTACAGAGCTGACGGCTGCTGCTGTGT	670
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	671	GCCGTGGCTGTGCCCCGCTGCAAGGGGCCACTCCCCACTGCTGCTGCTGCTGCTGCTG	730
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	731	GCTGCCGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTG	790
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	791	AGTGGCATCTGTGAGCTGCACTGCCAGGCCCTGGTCACTACCAACAGACACAGCTTTGAG	850
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	851	TCCATGCCCAATCCCAGGGCCGGTATACATTCCGGCCCGCAGCTGTGTGCTGCTGCC	910
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	911	TACAACCTACCTTTCTACGGAGCTGGGATCCTGCACCCCTCGTCTGCCCTGCACCAACCA	970
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	971	GAGGTGACAGCAGAGGATGCAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1030
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1031	GTGTGCTATGGTCTGGGCGATGGAGCACTTGCAGAGGTGGGGCACTTACCAGTGGCAAT	1090
QY	361	IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1091	ATCCAGAGTTTGTGCTGCTGCAAGAAGATCTTTGGGAGCTGGCAITTTCTGCCCGGAGCC	1150
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1151	TTTGATGGGAGCCAGCCCTCCAACACTGCCCGCTCCAGCCAGAGAGCTTCCAAGTGT	1210
QY	401	GluThrLeuGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1211	GAGACTCTGGAGAGATCACAGTTTACCTATACATCTCAGCATGGCCGACAGCCCTGCT	1270
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440







QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 Db 721 GCTGCGGCTGCACGGGCCCAAGCACTCTGACTGCCTGCCCTCCACTTCAACCAC 780  
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 Db 781 AGTGGCATCTGTGAGCTGCACCTGCGCAGCCCTGGTCACCTAACACACAGACAGCTTTGAG 840  
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 Db 841 TCCATGCCCAATCCCGAGGCGGTATACATTCGGGCGCAGCTGTGTGACTGCGCTGCC 900  
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 Db 901 TACAACATCTTCTACGACGCTGGGATCTCGACCTCGTGGCCCTCGCACACACAA 960  
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 Db 961 GAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCCA 1020  
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 Db 1021 GTGTGCTATGGTCTGGCATGGACACTTGCAGAGGTGAGGCGAGTTACCATGCCAAT 1080  
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 Db 1081 ATCCAGGAGTTTGTGCTGCAGAGAGATCTTTGGGAGCTGGCATTTCTGCCGGAGAGC 1140  
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluLeuGlnValPhe 400  
 Db 1141 TTTGATGGGAGCCAGCTCCAAACACTGCCCGCTCCAGCCAGAGCAGCTTCAAGTGT 1200  
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro 420  
 Db 1201 GAGACTCTGAGAGATCACAGGTACCTATACATCTCAGCATGGCCGACAGCTGCCT 1260  
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 Db 1261 GACCTCAGCGCTTCCAGAACCTTGCAAGTAATCCGGGGACGAATCTGCACAATGGCGC 1320  
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 Db 1321 TACTCGCTGACCTTCCAGGGCTGGGCATCAGCTGGCTGGGCTGGCGTCACTGAGGAA 1380  
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480  
 Db 1381 CTGGCAGTGGACTGSCCTCATCCACCAATACACCCACTCTGCTCTGCACACGGTG 1440  
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 Db 1441 CCCTGGGACACGCTCTTTCGGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCA 1500  
 QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 Db 1501 GAGGACAGTGTGTGGCGAGGGCTGGCTGCCACGCTGTGGCCCGGAGGGCACTGC 1560  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 Db 1561 TGGGCTCCAGGCGCCACCACTGTGTCACTGCAGCACTTCTCTGGGGCCGAGGAGTC 1620  
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
 Db 1621 GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680  
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 Db 1681 TTGCCCTGCCACCTTGTGAGTGTGAGCCCAAGATGGCTAGTGCCTGTTTGGACCGGAG 1740  
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
 Db 1741 GCTGACCACTGTGTGCCCTGTGCCACTATAGGACCTTCCCTCTGCGTGGCCGCTGC 1800  
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620

Db 1801 CCCAGCGGTGTAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860  
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspGly 640  
 Db 1861 GCGCATGCCAGCCTTGGCCCATCAACTGCACCACTCTCTGTGTGACCTGGATGACAAG 1920  
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
 Db 1921 GCGTCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCC 1959  
 RESULT 6  
 AX201817 3768 bp DNA linear PAT 30-AUG-2001  
 LOCUS Sequence 1 from Patent WO0153463.  
 DEFINITION AX201817  
 ACCESSION AX201817  
 VERSION AX201817.1 GI:15391666  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Cheever, M.A. and Hand-Zimmermann, S.  
 TITLE Compounds and methods for prevention and treatment of her-2/ neu  
 associated malignancies  
 JOURNAL Patent: WO 0153463-A 1 26-JUL-2001;  
 CORIXA CORPORATION (US)  
 FEATURES  
 source Location/Qualifiers  
 1..3768  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 1..3768  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC60126.1"  
 /db\_xref="GI:15391667"  
 /translation="MELAAALCRWGLLLALLPPGAASQVCTGTMKRLRPAASPETHLD  
 MURHLVQGVQVQGNLELYLPTNASLFLDIOEVQGVVLAHNOVRQVPLQRLRV  
 RGTOLFEDNYALAVLDNDGPLANTPTVTGASPGGLRELQRLRSITETLKGVLQIRNPQ  
 LCYQITLWKDIFHKNQALALHLIDITNRSRACHPCSPMCKGRCWGESSEDCQSLRT  
 VTAGGACRKGPLPTDCCHEQCAAGCTGPKHSDCLACLHNHSGICELHCPALVTNT  
 DTFSMPNPEGRYTFGACVCTACPNYLSDVGSCITVCLPHNQEVTAEDGDTQRCCK  
 SKPCRVCYGMEHLREVRVTSANIQFAGCKITFGSLAFLEPFSDGDPASNTAPL  
 OPEQLQVFTLEETGYLYISAWPDSLPSVFNQLOVIRGLTHNGAYSLTLOGLGI  
 SWLGRSLRELGSGLALIHNTHLICFVHTVPMDLFRNPQALHTNANRPECVCEG  
 LACHOLCARGHCWGPGTQCVCNCSQFLGQECVEECRVLQGLPREVYNARHCLPCHPE  
 CQPNGSVTCFGEADQCACAHYKDPFPCVACPSGVKPDLSYMPFIWPFDEEGACQ  
 PCPINTHSCVDLDDKCPAEQASPLTSIIISAVVGLLVGLVGVVGLIIRKROKI  
 RKYTMRRLLQETELVEPLTPSGAMPNQAQMRILKETELRKVKVYLSGAFGTVYKGI  
 PDGENVKIPVAIKVLENTSPKANKEILDEAYVMAGVGSYVSRLLGICLTSTVOLVT  
 OLMPYGCILLDHVNRGLGSODLLNMCQIAKMSYLEDVRLVHRDLAARNVLVKSP  
 NIWKTDTGLARLLDIDETEHADGKVPKIKWMALESILRRRTHQSDVWSVGYVWE  
 LMTFGAKPDGIPAREIDPLEKBERLPOPICTIDVYIMVKNWIDSECRPREL  
 VSEFSARMARDPQRFVYIQNEDLPASPLDSTFYRSLLEDDMDGLVDABEYLVPOQGF  
 DLPDAPAGAGGMVHHRRSSSTSGSGDLTLGLESEEAAPRSLAPSGAGSDVDFG  
 DLGMAAGKQLSLPTDPSPLQRYSEDPTVPLPSETDGYVAPLTGSPQEVYNQPDVR  
 POPPSRPGPLPAARPAGATLERPKTLPKNGVYKDFAFGAGVENPEYLTPOGGA  
 PQHPPPAFSPAFDNLTYWDQPPRGPAPPTFKGTPTAENPEYLGLDVPV"  
 BASE COUNT 759 a 1171 c 1119 g 719 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,84e-226 Length: 3768  
 Score: 3628.00 Matches: 653  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 SE03 (1-653) x AX201817 (1-3768)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20  
 Db 1 ATGGAGCTGGCGGCTTGTGGCGTGGGGCTCTCTCGCCCTTGTGCCCGGAGCC 60  
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
 Db 61 GCGAGCACCCAAAGTGTGACCGCGCACACATGAGCTGCGGCTCCCTGCGAGTCCCGAG 120  
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
 Db 121 ACCACCTGGAGATGCTCCGCCACTCTACACAGGCTGCCAGGTGGTGACGGAAACCTG 180  
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
 Db 181 GAATCACTACCTGCTCCGCCAACATGCCAGCGCTGCTTCCTGCAGGATATCCAGAGGTG 240  
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
 Db 241 CAGGGCTACGTGCTCATGCTCACAACCAAGTGAAGGAGGCTCTTGCATCCAGCGAAACCCGAG 300  
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
 Db 301 ATTTGGCAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGGTGCTAGACAATGGA 360  
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
 Db 361 GACCGCTGAACAATACCAACCCCTGTACAGGGGCTTCCCGCAGGAGGCTGCGGAGCTG 420  
 QY 141 GlnLeuArgSerLeuThrGluLeuLysGlyValLeuIleGlnArgAsnProGln 160  
 Db 421 CAGTTGCAAGCCCTCAGAGATCTTGAAGAGGAGGCTCTTGCATCCAGCGAAACCCGAG 480  
 QY 161 LeuCysTyrGlnAspThrIleLeuTriPlysAspIlePheHisLysAsnAsnGlnLeuAla 180  
 Db 481 CTCGTACACAGGACAGATTTTGTGAAGAGACATCTCCACAGAACCAACCACTGGCT 540  
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
 Db 541 CTCACATGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGCTTCTCCGATGTGAAG 600  
 QY 201 GlySerArgCysTrpGlyGluSerGlnAspCysGlnSerLeuThrArgValCys 220  
 Db 601 GGCTCCCGCTCTGGGAGAGAGTCTGAGAGATTGTGAGCCCTGACGCGCACCTGTCTGT 660  
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240  
 Db 661 GCCGGTGGCTGTCCCGCTGCAAGGGCCACTGCCACTGACTGCTGCCATGACAGCTGT 720  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 Db 721 GCTCCCGCTGCACGGGCCCCAAGCACTCTGACTGCGCTGGCTGCCCTCCACTTCAAGCAC 780  
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 Db 781 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGTGTCACCTACAAACACAGACACCTTTGAG 840  
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 Db 841 TCCATGCCAATCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCGCTGTCC 900  
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 Db 901 TACAACCTACCTTTCTACGGAGGTGGATCTCTGACCCCTCGCTCTGCCCGCTGCACAAACAA 960  
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 Db 961 GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAGTGCACAGACCCCTGTGCCGA 1020  
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 Db 1021 GTGTGCTATGCTGTGGCATGGACACTTTCGAGAGGTGAGGCGAGTTACAGTGCCTAAT 1080  
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380

Db 1081 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC 1140  
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
 Db 1141 TTTGATGGGACCCAGCCTCCAACTGCCCCGCTCCAGCCAGAGAGACTCCAAGTGT 1200  
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420  
 Db 1201 GAGACTCTGGAGAGATCACAGTTTACCTATACATCTCAGCATGCGCGGACAGCCTGCCT 1260  
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
 Db 1261 GACCTCAGCTGTCTCCAGAACCTCAAGTAATCGGGGACGAATTTCTGCAATATGGCGCC 1320  
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 Db 1321 TACTCGCTGACCTGCAAGGGCTGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGAA 1380  
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480  
 Db 1381 CTGGCAGTGGACTGGCCCTCATCCACATTAACACCCACCTCTGCTTCTGTCACACGGTG 1440  
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 Db 1441 CCCTGGGACCAAGCTCTTTGGGAACCCGACCAAGCTCTGTCTCCACACTGCGCAACCGGCCA 1500  
 QY 501 GluAspGluCysValGlyGluLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 Db 1501 GAGACAGAGTGTGGCGAGGGCTGGCCCTGCCACAGCTGTGCGCCGAGGACACTGC 1560  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 Db 1561 TGGGCTCCAGGGCCCAACCTGCTCAACTGACAGCAGTTCCTTCGSGGGCCAGAGTGC 1620  
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
 Db 1621 GTGAGGAATGCCAGTACTGACGGGCTGCCAGGGAGTATGTGAATGCCAGGCACTGT 1680  
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
 Db 1681 TTGCGGTGCCACCTGAGTGTGACGCCAGAAATGGCTCAGTGCAGTCTGTTTGGACCGGAG 1740  
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
 Db 1741 GCTCACCAGTGTGGCTGTGCCACTATAAGGACCTCCCTTCTGCTGCGCCGCTGC 1800  
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
 Db 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860  
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640  
 Db 1861 GGCCATGCCAGCCTTGCCCATCACTGACCCACCTCTCTGTGTGGAGCTGGATGACAAG 1920  
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
 Db 1921 GGCTGCCCGCGAGCAGAGAGCCAGCCCTCTGAGCTCC 1959

RESULT 7

AX380923

LOCUS

DEFINITION

AX380923

ACCESSION

AX380923.1

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

AX380923 3768 bp DNA linear PAT 18-MAR-2002  
 Sequence 9 from Patent WO0212341.

GI:19575767

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Cheever, M. A. and Gheysen, D.

Her-2/neu fusion proteins

JOURNAL Patent: WO 0212341-A 9 14-FEB-2002;  
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)  
FEATURES  
Location/Qualifiers  
1..3768  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1..3768  
/feature="human Her-2/neu protein"  
/codon\_start=1  
/protein\_id="CAD28392.1"  
/db\_xref="GI:19573768"  
/translation="MELAAALRWGLLLALLPGLAASQVCTGTDMLRLPASPETHLD  
MLRHLQGVQVQVQVLELTYPTNASLQDIQEVQGYVLIHNOVQVPLQRLIRY  
RGTOLFDENYLDNDGDLNNTPTVTGASPGRLQLRLSLTEILAGVYLIQRPQ  
LCYQTLWKDITPHKNQALATLIDNRSRACHPCSPMKGSRGSESDQSQSLTRT  
VCAGCARCKGLPDKDCHQCAAGCTGPKHSDCLACLFHNSHGICEHLRCPALVTNT  
DTFESMNPREGRYTEGACVTPACPYNYLSTDVGSCTLVCPHLNQEVTAEDGTQCEKC  
SKPCARVCIQGLMEHLREVRNVTASIQEFAGCKKIFGLAFLPESFGDDPASNTAPL  
QPEQVVFETLEEITGYLIYISAWPDSLPDLVSFQNLQVIRGLHNGAYSLSLQGLGI  
SWGLRSLEIGSLGALIHNTHTLCFVHTVWDQLFRNPHQALLHNRDEDECYGE  
LACHOLCARHGWGPQTCVNCQFQVLRQECVRECRVLOGLPREYVNRHCLPCPE  
COPQNSVTCGPEADOCVACAHYKDPFCVACRPSGVKPDLSYMPJWKFDEEGACQ  
PCPINCTHSCVDLDKGCPCAPQASPLISISAVVGLLVLVGVVFGVGLIKRRQKI  
RYKTHNQLQETELVEPLTPSGAMPNQAMRILKTELRKVKVLGSGAFVYKGIWI  
PDGENVKIPVAIKLRENTSPKANKEILDEAYVMAGVSPYVSRLLGLCLSTVOLVT  
OLMPYGLLDHVRNRRGLGSODLLNMCQIAKMSYLEDVRLVHRDLAARNVLKSP  
NHVITDTEGLARLLDIDETEVHADGKVPKIKMALESILRRFTHOSDVWSYGVTVWE  
LMTFCAPYDGIPIAREIDLEKGERLPQPPCTIDVYHIMVKHMDISECRPREL  
VSEFRMARDQRFVYIQNEDLGPASPLDSTFYRSLLDDDDMDGLDVAEYLVPQGF  
FLCPAPAGAGMVHRRSSRSGGDLTLGLEPSEEEAPRSLAPSEAGSDVFDG  
DLMGAARGLASLPDPSPQRYSEDTPVLPSETDGYVAPLTCSPQPEYVNPQDVR  
POPPSPREGPLPAAPAGATLERKTLSPGKNGVVKDVFAGGAVENPEYLTPOGGA  
POPHPPAFSPAFDNLVYDQDPPPERGAPPTFKTPTPAENPEYLGDPV"  
1..1959  
/misc\_feature  
/feature="extracellular domain (ECD) of human Her-2/neu"  
2026..3765  
/misc\_feature  
/feature="intracellular domain (ICD) of human Her-2/neu"  
2968..3765  
/misc\_feature  
/feature="phosphorylation domain (PD) of human Her-2/neu"  
2968..3144  
/misc\_feature  
/feature="preferred portion of the phosphorylation domain  
(delta PD) of human Her-2/neu"  
759 a 1171 c 1119 g 719 t

BASE COUNT 759 a 1171 c 1119 g 719 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,84e-226 Length: 3768  
Score: 3628.00 Matches: 653  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

SEQ3 (1-653) x AX380923 (1-3768)

QY 1 MetGluLeuAlaLeuCysArgTrrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla 20  
|||||  
Db 1 ATGGAGCTGGCGGCTTGTGGCGCTGGGGGCTCTCTCGCCCTCTTGGCCCCCGAGGCC 60  
|||||  
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
|||||  
Db 61 GCGAGCACCCAAAGTGTGCACCGGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 120  
|||||  
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60  
|||||  
Db 121 ACCCACTGGACATCTCCGCACCTCTACCAAGGCTGCCAGGCTGGTGCAGGGAACCTG 180  
|||||  
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
|||||  
Db 181 GAACATCACTACTCTGCCACCCACCAATGCCAGCGTGTCTTCCGTCAGGATATCCAGAGGTG 240  
|||||  
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
|||||

|||||  
241 CAGGCTACGTGCTCATGCTCACAACTAGTAGGAGTCCCACTGCAGAGGCTGCGG 300  
|||||  
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
|||||  
Db 301 ATTGTGCGGAGCACCAGCTCTTTGAGGACAACTATGCTTGGCCGTGCTAGACAAATGA 360  
|||||  
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
|||||  
Db 361 GACCGCTGAACAATACCACCCCTCTCACAGGGGCTCCCCAGGAGGCTGCGGGAGCTG 420  
|||||  
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160  
|||||  
Db 421 CAGCTTCGAAGCCTCACAGATCTTGAAGAGAGGGGTCTTATCCAGCGGAACCCAG 480  
|||||  
QY 161 LeuCysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnAsnGlnLeuAla 180  
|||||  
Db 481 CTCTGCTACAGGACACCATTTTGTGAAGAGACATCTTCCACAAAGAACACCGCTGGCT 540  
|||||  
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
|||||  
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 600  
|||||  
QY 201 GlySerArgCysTrrpGlySerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
|||||  
Db 601 GGCTCCCGCTGCTGGGGAGAGAGTTCTGAGGATTTCTCAGAGCTGACGCGCACTCTCTCT 660  
|||||  
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
|||||  
Db 661 GCCGTGGCTGTGCCCGCTGCAAGGGGCACTGCCCACTGACTGTGCTCATGAGCAGTGT 720  
|||||  
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
|||||  
Db 721 GCTGCCGCTGCACGGGCCCCAAGACACTCTGACTGCTGCTGGCTGCTTCCACTTCAACCC 780  
|||||  
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
|||||  
Db 781 AGTGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTACACACACAGACACGTTTGAG 840  
|||||  
QY 281 SerMetProAsnProGluGlyArgTrrpPheGlyAlaSerCysValThrAlaCysPro 300  
|||||  
Db 841 TCCATGCCCAATCCCAGGCGGTATACATTCGGCGGCGAGTGTGTGACTGCTGCTGCC 900  
|||||  
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
|||||  
Db 901 TACAACTACTTCTACGAGCTGGGATCTGCACCCCTCGTGTGCCCTTGCACAAACCA 960  
|||||  
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
|||||  
Db 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAAGTGCAGCAAGCCCTGTGCCCA 1020  
|||||  
QY 341 ValCysTyrGlyLeuGlyMetClnHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
|||||  
Db 1021 GTGTGCTATGGTCTGGGATGGAGCACTTGGAGAGGTTGGAGGCGAGTTACCAAGTCCCAAT 1080  
|||||  
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
|||||  
Db 1081 ATCCAGGAGTTGTGGCTGCAAGAAGATCTTTGGAGGCTGGCATTTCTGCCGAGAGC 1140  
|||||  
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
|||||  
Db 1141 TTTGATGGGACCCAGCCTCCAAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT 1200  
|||||  
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrrpProAspSerLeuPro 420  
|||||  
Db 1201 GAGACTCTGSAAGAGATCACAGGTTCATCATATACATCTCAGCATGCCCGGAGCAGCTGCT 1260  
|||||  
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
|||||  
Db 1261 GACCTCAGCGTCTTCCAGAACCTGCAAGTAAATCCGGGGAGCAAAATTTCTGCACAAATGGCGCC 1320  
|||||  
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
|||||



Db	1581	TTGCCGTGGCCACCTCGAGTGTACGCCGCCAGAAATGGCTACGTGACCTGTTTGGTGGACCGGAG	174
Qy	581	AlaAspGlnCysValAlaAlaCysAlaHisTyrLysAspProProPheCysValAlaAlaArgCys	600
Db	1741	GCTGACCAGTGTGTGGCTGTGTGCCACTATAAGGACCCCTCCCTTCTCGTGGCCCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCCAGCGGTGTGAACCTGACCTCTCCTACATGCCCATCTCGGAAGTTTCCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCCATGCCAGCCCTGCCCATCACTGACCCACCTCCTGTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCCGAGCAGAGCCAGCCCTCTGACGTCC	1959
RESULT	9		
AX465456			
LOCUS	AX465456	3768 bp	DNA linear PAT 16-JUL-2002
DEFINITION	Sequence 1 from Patent WO02113847.		
ACCESSION	AX465456		
VERSION	AX465456.1	GI:21899819	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE			
AUTHORS	Galger, A., Cheever, M.A. and Hand-Zimmermann, S.		
TITLE	Methods for diagnosis and therapy of hematological and virus-associated malignancies		
JOURNAL	Patent: WO 0213847-A 1 21-FEB-2002;		
	CORIXA CORPORATION (US)		
FEATURES			
source	location/Qualifiers		
	1..3768		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
CDS	1..3768		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD42363.1"		
	/db_xref="GI:21899820"		
	/translations="MELAAICRWGLLIALLPGGAASQVCTGTMKLRLPASPETHLD MLRHLQGCQVQVGNLELYLPINASLSFLQDIEQVGYVLIHQNVRQVPLQRRLV RGYDFLMDKTLAVLQNDPLNPTVPGASGLRELQRLSLEILKGGLVLIORNPQ LCYQDTILKDTLIPHKNNQALTLIDNRSRACHPCSPMKSCRWGSSDDCSQVLT VCAGARCKRGLPTDCCHPQCAAGCTGPKHSDCLACLHPNHSIGICELHCPALVYTN DTFESNPNBPTVFAGSCVTACPYNVLSTDVGCTLVCPHNOEYTAEDGTQCKEC SKPCARVCYGLGMEHLREVRVAVTANIOEPAGCKLFGSLAFILPESFDGDPASNTAPL QPQLOVFTELREITGYLYTISANPDSLPLDSVFQNLQVIRGRILHNGAYSLTLQGLGI SWLGRSLRELGSLGALIHINTHLCFVYVPMQDLFRNPHQALLHTANRDECVGEG LACHQLCARHGCHPGPTQVCNCSQFLRGQCEYEECVRLQGLRPYVNAHCLPCHPE CQPNQSGTFCGEAPQACVAAHYKDPFCVCARCPGKVPDLSPYIMKFPDDEGACQ PCPINCTHSCVDLDDKCPQCAORASPLTSTISAVVGILLVVVLGVVFGILIKRQOKI RKYTMRLQETELVLEPLTPSGAMPNOAMRILKETELRKVKYVLSGAGFTVYKGIWI PDGENVKIPVAIKVLENTSPKANKELDEAYVMAGVSPYSRLIGLICLTSTVQLVT QLMYPGCLLDVRENRGQSOLLNCMOIAGNSYLEDVRLVLRHDLAARNVLYKSP NHVKITDFGLARLLIDETGYHADGGKVPYIKWMALESILRRRTHQSDVMSYGVTVWE LMTFGAPYDGIPIARETIDLEKGERLPQPICTIDVYIMVKWCMIDSECRPRFREL VSTFESMARDPQRFVYIIONEDLGPASPLDSTFYRSLLDEDDMGDLVDAREYLYPQOFG FCPDPAFGAGMVVHHRSSSTSGGGDLTLGLEPSEEAAPSPPLAPSGAGSDVPDG DLGMAKAGLQSLDHPDTPSLQRYSEDTVPVLPSETDGIYVAPLTSQPPEYVNOPDVR PQPSREGPLPAARFAGATLERPKTSLSPGKNVVKDVFAGFAGVENPEYLLTPQGGAA PQPHPPAPSPAFDNLVYMDQDPPERGAAPPSTPKGTPTAENPEYLLGLDVPV"		
BASE COUNT	759 a	1171 c	1119 g 719 t
ORIGIN			
Alignment Scores:			
Pred. No.:	2.84e-226	Length:	3768
Score:	3628.00	Matches:	653

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
SEQ3 (1-653) x AX465456 (1-3768)  
Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla 20  
Db 1 ATGGAGCTGGGGCGCTGTGGCGCTGGGGCTCCTCGCCCTCTTGGCCCGGAGGCC 60  
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40  
Db 61 CGAGACCCCAAGGTGACCCCGGACAGACATGAAGCTGGCGCTCCCTGCCAGTCCCGAG 120  
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlyAsnLeu 60  
Db 121 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGACGGGAACCTG 180  
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80  
Db 181 GAACCTACCTACCTGCCACCAATGCCAGCTGTCTCTCTGCGAGATATCCAGGAGGTG 240  
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100  
Db 241 CAGGGCTACGTGCTCATCGCTCACAACTGAGGAGGCTGCCACTGCAGAGGCTGCGG 300  
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120  
Db 301 ATTGTGGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGGTGCTAGACAAATGA 360  
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140  
Db 361 GACCGCTGAACAATACCAACCCCTGTACAGGGGCTCCCGAGGAGGCTCGGGAGCTG 420  
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160  
Db 421 CAGCTTGCAGCCCTCAGAGATCTTGAAGAGGAGGGGTCTTGAACGCGGAACCCCGAG 480  
Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180  
Db 481 CTCTGCTACCAAGACAGATTTTGGNAGGACATCTCCACAGACAAACCACTGGCT 540  
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200  
Db 541 CTCACCTGTATGATACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 600  
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
Db 601 GGCTCCCGCTCTGGGAGAGAGTTCTGAGGATTGTGAGAGCCTGACGCGCAGCTGTCTGT 660  
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
Db 661 GCCGGTGGCTGTGCCCGCTGCAAGGGCCACTGCCACTGACTGCTGCCATGACAGCTGT 720  
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
Db 721 GCTCCGGCTGCACGGGCCCCAAGCACTCTGACTGCGCTGGCGCTGCCCTCAACCCAC 780  
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
Db 781 AGTGGCATCTGTAGCTGCACTGCCCGCCCTGTATACATTCGCCCGCAGCTGTGTGACTG 840  
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
Db 841 TCCATGCCCAATCCCGAGGCGCGGTATACATTCGCCCGCAGCTGTGTGACTGCTGCC 900  
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
Db 901 TACAACCTACCTTTCTACGGAGCTGGGATCTCTGCACCTCGCTCGCTGCCCAACCAA 960  
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGlnLysCysSerLysProCysAlaArg 340

Db 961 GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAAGTGACAGCAAGCCCTGTGCCCGA 1020  
Qy 341 ValCysTyrGlyLeuGlyMetGlnHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
Db 1021 GTGTGCTATGCTTGGGCATGGAGCACTTGGAGAGGTGAGGGCAGTGTACCAAGTCCCAAT 1080  
Qy 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
Db 1081 ATCCAGGAGTTGTGCTGCTCAAGAAGATCTTTGGGAGCCTTGGCATTTCTGCCCGAGAGC 1140  
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400  
Db 1141 TTTGATGGGACCCAGCCTCCAACTGCCCGCTCCAGCCAGAGCACTCCAAGTGT 1200  
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420  
Db 1201 GAGACTGTGGAAGAGATCACAGGTACCTATATCATCTCAGCATGCCCGGACAGCCTGGCT 1260  
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440  
Db 1261 GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGACGAATTTCTGCACAATGGCGCC 1320  
Qy 441 TyrSerLeuThrLeuGlnGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
Db 1321 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTCGCTCACTGAGGGAA 1380  
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480  
Db 1381 CTGGCAGTGGACTGGCCCTCATCCACATTAACACCCACCTCTGCTTGTGCACACGGTG 1440  
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
Db 1441 CCTTGGGACCAAGCTCTTTCCGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500  
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
Db 1501 GAGACAGTGTGGGCGAGGGCTGGCTGCCACCAAGCTGTGGCGCCGAGGCACTGC 1560  
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
Db 1561 TGGGGTCCAGGGCCCCCAGCTGTGCTCACTGCAGCAGATTCCTTCGGGGCCAGGAGTGC 1620  
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
Db 1621 GTGAGGAATGCCAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACCTGT 1680  
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
Db 1681 TTGCGGTGCCACCTGAGTGTGACGCCCGAGAAATGGCTCAGTGACCTGTTTGGACCGGAG 1740  
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
Db 1741 GCTACCAAGTGTGGCTGTGCCCACTATTAAGACCCCTCCCTTCTGCTGGCCCGCTGC 1800  
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
Db 1801 CCCAGCGTGTGAACCTGACCTCTCTCTACATGCCCATCTGCAAGTGTCCAGATGAGGAG 1860  
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640  
Db 1861 GGCCGATGCCAGCCTGCCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 1920  
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
Db 1921 GGCTGCCCGCCGAGCAGAGAGCCGCTCTGAGCTCC 1959  
RESULT 10  
AX467229  
LOCUS AX467229 3768 bp DNA linear PAT 16-JUL-2002  
DEFINITION Sequence 3 from Patent WO0234287.  
ACCESSION AX467229  
VERSION AX467229.1 GI:21900511  
KEYWORDS

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Beier, A.M., Gautam, A. and Mouritsen, S.R.
TITLE       Novel therapeutic vaccine formulations
JOURNAL     Patent: WO 0234287-A 3 02-MAY-2002;
            Pharmexa A/S (DK)
FEATURES
  Source     1...3768
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
  CDS        1..3768
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD2452.1"
            /db_xref="GI:21900512"
            /translation="MELAAALCRWGLLLALLPFGAASQTQVCTGTDMLKRLPASPEETHLD
            MRHLVQGVVOCNLELTYPNTWASLSFLQDIOEVGYVLIHNOVROVPLQRLRV
            RGTOLFEDNVALVLDNGDPLNNTPTVTGSPGGLRELQRLSLTEILKGGVLIQRPQ
            LCYQDTILKWDIFHKNNQALALITDNRSRACHPCSPMKGSRGWSGESDCQSLTRT
            VCAGGACRGPLEPTDCCHEQCAAGCTGPKHSDCLACLFNHSIGICELHCPALTYNT
            DTFSMPNPGRYTFGASCTFACPNYLSTDVGSCTLVPLHNOEVAEDGTQRCRC
            SKPCRVYIGIMEHLREAVTSANIQEFAGCKIFGSLAFLPESDGDYASLTQGLGI
            QPEOLVFTELEITGYLISAWPDSLPLDSVFONLOVIRGRILHNGAYSILTQGLGI
            SWGLSLRELGSGLIHHNTHLCFVHTVPMQDLFRNPHQALLHTANRDEDECYGG
            LACHQLCARHCWGPPTQCNCSQFLRGQECVEECRVLQGLPREYVNRHCLCPHE
            CQPQNSVTCFPGPADQCACAHYKDPFCVARGPSGVKPLDSYMPLWKPFDEGAQO
            PCPIFNHSCVDLDDKCPAPASPLTSIVSAVGLLVVLVGVGAFGTIKRQQRKI
            RKTMYRLLEDELVEPLTSPGAMPNQMRILKETELRKVKVLGSGAFGTVYKGIWI
            PDGENVPAIVAKILRENTSPKANKEILDEAYVWAGVSPVSRLLGTLTSTVOLVT
            QLMFYGLDVRNRLGSLQDLNMCMIAXMGMSVLEQVRLVHRLDAAARNVLKSP
            NHVITDFGLARLIDIDETEHADGKVPFKWMALESILRRRFTHQSDVMSYGYTVE
            LWFSGARPDGPAREIPDLLEGRLPDPPTCYTIDVIMVKNWCMIDSECRPFREL
            VSEFRWARDQPREVIONEDLGPASLDSTFYRSLLDDDDMDGLVDAEYLYPQGF
            FCPDPAAGAGMWHRRSSTRSGGDLTLGLEPSEEAERPSPLAPSEAGAGSDVFG
            DLGMAAGAGLSLPTDPSPLQRYSEDTPVLPSETDGYVAPLTCSPQPEVNVQPDYR
            POPPSREGPLPAAPAGATLERAKTLSPGKNVGVKDFVAFGGAVENPEYLTPOGGAA
            POPHPPAFSPAFDNLVWDQDPPERCAPSTFKGTPTAENPEVLGLDVPV"
BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN
Alignment Scores:
Pred. No.:      2.84e-226      Length:      3768
Score:          3628.00        Matches:      653
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:     100.00%      Indels:        0
DB:              6            Gaps:         0
SEQ3 (1-653) x AX467229 (1-3768)
Qy      1 MetGluLeuAlaLaLeuCysArgTrrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla 20
Db      1 ATGGAGCTGGGGCTTGTGCGCTGGGGCTCTCTCGCCCTCTTGGCCCGGAGCC 60
Qy      21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db      61 GCGAGACCCCAAGTGTGCACCGGCACAGATGAAGCTGGGGCTCCCTGCGAGTCCCGGAG 120
Qy      41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAspLeu 60
Db      121 ACCACCTGGACATGCTCCGCCACCTCTACCGAGGCTGCCAGGTGGTGCAGGGAACCTG 180
Qy      61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnGlnVal 80
Db      181 GAATCACTACCTGCTGCCCAATGCCAGCTGCTCTTCTGTCAGGATATCCAGGAGGTG 240
Qy      81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      241 CAGGGCTACGTGCTCATCGCTCACACCAAGTGCAGGAGGTGCCACTGCCAGTCTGCGG 300

```

```

101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      301 ATTTGGCAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTAGACAAATGGA 360
Qy      121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db      361 GACCGCTGAACAATAACCACTCTGTACAGGGGCTCCCCAGGAGGCTCGGGGAGCTG 420
Qy      141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db      421 CAGCTTCGACCCCTCACAGAGATCTTGAAGAGAGGGGTCTTGATCCAGCGGAACCCAG 480
Qy      161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
Db      481 CTCTGCTACCAAGACAGATTTTGTGGAAGAGACATCTTCCACAAAGAACCAACGCTGCT 540
Qy      181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 600
Qy      201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      601 GGCTCCGCTCTCGGGAGAGATTTCTGAGGATTTCTGAGGCTGACGCGCTCTCTGT 660
Qy      221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db      661 GCCGTGGCTGTGCCCTGCAAGGGCCACTGCCCACTGACTGCTGCATGAGCAGGTGT 720
Qy      241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db      721 GCTCGGCTGCACGGGCCCCAAGCACCTCTGACTGCTGCTGCTGCTCCACTTCAACCAC 780
Qy      261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db      781 AGTGGCATCTGTGAGCTGACTGCCACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy      281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db      841 TCCATGCCAATCCGAGGCGGTATACATTCGGCGCCAGCTGTGTGCTGCTGCTGCTGCT 900
Qy      301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db      901 TACAACCTACCTTTCTAGCGAGCTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy      321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db      961 GAGGTGACAGCAGGAGATGGAACACAGCGGTGTGAGAAAGTGCAGCAAGCCCTGTGCTG 1020
Qy      341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db      1021 GTGTGCTATGCTTGGGCATGGAGCCTTGGAGAGGTGAGGGCAGTTTACCAGTGCCTAAT 1080
Qy      361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db      1081 ATCCAGGAGTTTGTGCTGCAAGAAGATCTTTGGAGCTTGGCATTTCTGCCGAGAGC 1140
Qy      381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db      1141 TTTGATGGGACCCAGGCTCCAAACACTGCCCGCTCCAGCCAGCAGCAGCTCCAAAGTGT 1200
Qy      401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrrpProAspSerLeuPro 420
Db      1201 GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGCGCGGACAGCTGCT 1260
Qy      421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db      1261 GACCTCAGCGCTTCCAGAACCTCAAGTAATCCGGGACGAAATTCGCACAAATGGCGCC 1320
Qy      441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db      1321 TACTCGTGCCTTGCAGGGCTGGCATCAGTGTGGTGGGCTGCGCTACTGAGGAA 1380

```



```
QY 461 LeuGlySerGlyLeuAlaLeuLeuHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 CTGGCAGTGGAGTGGCCCTCATCCACCAATAACACCCACTCTGCTTCGTGTCACAGCGT 1440
QY 481 ProTirAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCTGGGACCAAGCTCTTTCCGAAACCCACCAAGCTCTGCTCCACACTGCCACACCGGCCA 1500
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGGACGAGTGTGGSCGAGGCGCTGGCCCTGCCACAGCTGTGGCCCGCAGGCACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGGTCCAGGGCCCAACCAAGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 GTGGAGGAATGCCGAGTACTGCAAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCGGTGCCACCTGAGTGTGACCCCAAGATGGCTCAGTACCTGTTTGGACCGGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1741 GCTGACCAAGTGTGGGCTGTGCCACTATAAGACACCTCCCTTCTGCGTGGCCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleThrLysPheProAspGluGlu 620
Db 1801 CCCAGCGGTGAAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1861 GGCAGTCCAGCAGCTTGCCTCCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACA 1920
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCCCGAGCAGAGAGCCAGCCCTCTGACGTCC 1959

RESULT.11
LOCUS AX481438 3768 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 52 from Patent WO02055693.
ACCESSION AX481438
VERSION AX481438.1 GI:22316352
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
1 Kreutzer,R., Limmer,S., Rost,S. and Hadwiger,P.
AUTHORS Method for inhibiting the expression of a target gene
TITLE Patent: WO 02055693-A 52 18-JUL-2002;
JOURNAL Ribopharma AG (DE)
FEATURES
source
1. .3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN

Alignment Scores:
Pred. No.: 2,84e-226 Length: 3768
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Watch: 100.00% Indels: 0
DB: 6 Gaps: 0

SEQ3 (1-653) x AX481438 (1-3768)
```

```
QY 1 MetGluLeuAlaLeuCysArgTrrpGlyLeuLeuAlaLeuLeuProProGlyAla 20
Db 1 ATGGAGCTGGCGGCTGTGCTCCGCTGGGGCTCTCTCTCGCTTCTTGTCCCGGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCAGAGCAACCAAGTGTGCACCGGCACAGATGAAGCTGGGCTCCCTGGCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu 60
Db 121 ACCACCTGGACATGCTCCCGCCACTCTACAGGGCTGCCAGGTGTGTGTCAGGGAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 181 GAACTCACCTTACCTGGCCCAACCAATGCCAGCTGTCTCTCTGAGGATATCCAGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGTACCTGTCTATCGCTCACACCAAGTGAAGGAGGTCCCACTTCAGAGGCTCGCG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGGAGGACCCAGCTCTTTGAGGACAACATATGCCCTGGCCGTGCTAGACAAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCGCTGAACAATACCAACCCCTGTCCAGAGGGCTCCCAAGAGGCGCTCGGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTTGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCACGCGAACCACCCAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTrrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCTGTACCAAGACAGATTTGTGAAGGACATCTTCACAGAAACAACACAGCTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAAG 600
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGCTGTCTGGGAGAGATCTGAGGATTTGTCAGAGCTGTGAGCGCAGCTGTCTGT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCCGGTGGCTGTGCCGCTGCAAGGGGCCACTGCCCACTGCTGCTGCCATGAGCAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCCCTGCCCTCAACAC 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGGCATCTGTGAGCTGCACCTGCCACCGCTGCTACCTACACACAGACACAGCTTTGAG 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCCATGCCCAATCCCGAGGGCCGGTATACATTGGGCCAGCTGTGTGACTGCTGTCCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 TACAACCTACCTTTCTACGGAGTGGGATCTCGACCCCTGCTGCTGCCCTGCACAACCAA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 GAGGTGACAGCAGAGGATGAACACAGCGGTGTGAGAAGTGCAGCAGACCTGTGCGCGA 1020
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1021 GTGTGCTATGCTGGGCACTGGAGCACTTCGGAGAGGTGAGGGCAGTTACAGTGCACAT 1080
```



```
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
    |||
Db 1081 ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCAATTTCTGCGGAGACC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
    |||
Db 1141 TTTGATGGGGAGCCAGCCATCCAAACACTGCCCCGCTCCAGCCAGAGAGAGCTCCAAAGTGTTT 1200
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
    |||
Db 1201 GAGACTCTGGAAGATACAGAGTACCTATACATCTCAGCATGGCCGACAGCCCTGCGCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
    |||
Db 1261 GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATCTGCACAATGCGGCC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
    |||
Db 1321 TACTCGCTCAGCCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAA 1380
QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
    |||
Db 1381 CTGGCGAGTGGAGTGGCCCTCATCCACATACACCCACCTCTGCTTCGTGTCACACGGTG 1440
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
    |||
Db 1441 CCCTGGGACAGCTCTTTCGGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
    |||
Db 1501 GAGGAGAGTGTGTGGCGAGGGCTGGCCCTGCCACAGCTGTGGCCCGGAGGGCACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
    |||
Db 1561 TGGGGTCCAGGGCCACCAGCTGTGTCACTGACGACCACTTCTTCGGGGCCAGGAGTGC 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
    |||
Db 1621 GTGGAGGAATGGCGAGTACTGCGAGGGCTCCCGAGGGAGTATGTGAATGCCAGGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
    |||
Db 1681 TTGCGGTGACCCCTGAGTGTGAGCCCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
    |||
Db 1741 GCTGACCACTGTGTGGCTGTGCCACTATAGGACCTCCCTCTGCGTGGCCCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
    |||
Db 1801 CCCAGCGGTGTAAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
    |||
Db 1861 GCGCATGTCAGCCTTGCCCCATCAACTGCACCCACTCTGCTGTGGACCTGGATGACAAG 1920
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
    |||
Db 1921 GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCC 1959
```

RESULT 12

```
LOCUS AR080259 4473 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 1 from patent US 5968748.
ACCESSION AR080259
VERSION AR080259.1 GI:10006994
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4473)
AUTHORS Bennett,C.Frank., Lipton,A. and Witters,L.M.
TITLE Antisense oligonucleotide modulation of human HER-2 expression
```

```
JOURNAL Patent: US 5968748-A 1 19-OCT-1999;
FEATURES
Source 1. 4473
BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN
Alignment Scores:
Pred. No.: 3 45e-226 Length: 4473
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
SE03 (1-653) x AR080259 (1-4473)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
    |||
Db 175 ATGGAGCTGGCGCCTTGTGCCGTGGGGCTCTCTCGCCCTCTTGGCCCCCGGAGCC 234
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
    |||
Db 235 GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGCTCGGGCTCCCTGCCAGTCCCGAG 294
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
    |||
Db 295 ACCACCTCGACATGCTCGCCACCTCTACAGGGCTGCCAGGTGTGCAGGAAACCTG 354
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
    |||
Db 355 GAATCATCCTACCTGCGCCACCAATGCCAGCTGTCTTCTGCAGGATATCCAGGAGGTG 414
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
    |||
Db 415 CAGGGCTAGCTGCTCATGCTCACAAACCAAGTGAGGAGGTCCCATCGCAGAGGCTGCGG 474
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
    |||
Db 475 ATTGTGCGAGGCCACCGACTTTTGAGGACAACTATGCCCTGCCCTGTAGACAATGGA 534
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
    |||
Db 535 GACCCGCTGAACAATACCACCTCTGTACAGGGGGCTCCCGAGGAGGCTGCGGGAGCTG 594
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
    |||
Db 595 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAAACCCCGAG 654
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
    |||
Db 655 CTCTGCTACCAGACACGAGATTTTGGGAAGGACATCTTCCACAAGAACCAACCACTGGCT 714
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
    |||
Db 715 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAAG 774
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
    |||
Db 775 GGCTCCCGCTGCTGGGGAGAGAGTCTTGAGGATTTGTCAGAGCCTTGACGGCCACTGCTGT 834
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
    |||
Db 835 GCCGGTGGCTGTGCCCGCTGCAGGGGCCACTGCCCCACTGACTGTGCCATGAGCAGTGT 894
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
    |||
Db 895 GCTGCGGGCTGCACGGCCCCCAAGCACTCTGACTGCTGGCTGGCTGCCACTTCAACCCAC 954
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
    |||
Db 955 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTGCACCTACACACAGACACAGCTTTGAG 1014
```

```

QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 1015 TCCATGCCCAATCCCGAGGCGGTATACATTCGGCGCAGCTGTGTGACTGCTGTCC 1074

QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1075 TACAACATACCTTTCACGAGAGTGGGATCCCTGACCCCTCGTGTGCCCTGCACAAACAA 1134

QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1135 GAGGTGACAGCAGAGATGGACACAGCGGTGTGAGAAGTCAGCAAGCCCTGTGCCCGA 1194

QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1195 GTGTCTATGTCTGGCATGGAGCAGCTTGCAGAGGTGAGGGCAGTTTACAGTGCCCAAT 1254

QY 361 IleGlnGluPheAlaGlyCysAlaLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1255 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGGAGCCTGGCAATTTCTGCCGGAGAGC 1314

QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1315 TTTGATGGGACCCAGCCTCCAAACACTGCCCGCTCCAGCCAGACGACTCCAGTGT 1374

QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1375 GAGACTCTGGAGAGATCACAGGTATACCTATACATCTCAGCATGCGCGAGACGCTGCT 1434

QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1435 GACCTACGCTTCCAGAACCTGCAAGATATCCGGGAGCAATTTCTGCACAAATGGCGCC 1494

QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1495 TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTCGCTCTACTGAGGGAA 1554

QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1555 CTGGCAGTGACTGGCCCTCATCCACCATACACCCACCTCTGCTGCTGCACACGCTG 1614

QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1615 CCCTGGGACCACTCTTTTCGAAACCCGACCAAGCTCTGCTCCACACTGCCCAACCGGCA 1674

QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1675 GAGGACAGTGTGTGGCGAGGCGCTGGCCTGCCACCAAGCTGTGCGCCGAGGCGACTGC 1734

QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1735 TGGGTCAGGGCCACCCAGTGTGTCACCTGACGACAGTTCCTTCGGGGCCAGGAGTGC 1794

QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1795 GTGGAGGAATGCCGAGTACTCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1854

QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1855 TTGCCGTGCCACCTGTAGTGTGACCCAGCAATGGCTCAGTGACCTGTTTGGACCCGAG 1914

QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1915 CTGACCAAGTGTGGGCTGTGCCACTATTAAGACCCCTCCCTTCGCGTGGCCGCTGC 1974

QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1975 CCCAGCGTGTGAACCTGACCTCTCTACATGCCATCTCGAAGTTTCCAGATGAGGAG 2034

QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 2035 GGCAGATGCCAGCCTTGCCCATCAACTGCACCCACTCTGTGTGGACTTGATGACAAG 2094

QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653

```

```

Db 2095 GGCTGCCCGCGAGCAGAGAGCCAGCCCTCTGTGACGTCC 2133

RESULT 13
LOCUS ARI67390
DEFINITION Sequence 26 from patent US 6287569.
ACCESSION ARI67390
VERSION ARI67390.1 GI:17903168
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4473)
AUTHORS Kipps T.J. and Wu Y.
TITLE Vaccines with enhanced intracellular processing
JOURNAL Patent: US 6287569-A 26 11-SEP-2001;
FEATURES
    Location/Qualifiers
        source
            1..4473
                /organism="unknown"
BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN
Alignment Scores:
Pred. No.: 3,45e-226 Length: 4473
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

SEQ3 (1-653) x ARI67390 (1-4473)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProGlyAla 20
Db 175 ATGAGCTGGCGGCTTGTGCGCTGGGGCTCTCTCTCGCCCTTGTCCCGGAGCC 234

QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCGAGCACCAGTGTGCACCGGCACACATGAAGCTGCGCTCCCTGCCAGTCCCGAG 294

QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 295 ACCCACTGGACATGCTCCGCCACTCTTACAGGGCTGCCAGGTGGTGACAGGAACCTG 354

QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 355 GAATCCTACCTACCTGCCCAACCAATGCCAGCTGTCTTCTGCAGGATATCCAGGAGTG 414

QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGCTACGTGCTCATCGCTCACAACTGAGGAGGTCTTTCCTGCAGAGGTGCGG 474

QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 475 ATTTGGGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGCTGTAGACAATGA 534

QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGTGAACATACCCCTGTACAGGGGCTCCCGAGGAGGCTCGGGAGGCTG 594

QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 595 CAGCTTCAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGTATCCAGCGGAACCCCGAG 654

QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 655 CTCTGCTACCAAGGACAGATTTTGTGAAGAGACATCTTCCACAAGAACCAACAGCTGCT 714

QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 715 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCCGATGTGTAG 774

```

QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220  
 DB 775 GGCCTCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTCAGAGCCTGACGCGCACTGCTCT 834  
 QY 221 AlaGlyGlyCysAlaArgCysGlyProLeuProThrAspCysCysHisGluGlnCys 240  
 DB 835 GCGGGTGGCTGGCCCGCTGCAAGGGGCCACCTGCCACCTGACTGCTGCCATGAGCAGTGT 894  
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260  
 DB 895 GCTCGCGGCTGCACGGGCCCCAAGCACCTCTGACTGCTGCCCTGCCCTCCACTTCACAC 954  
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280  
 DB 955 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACCTACAAACACAGACAGCTTGAG 1014  
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300  
 DB 1015 TCCATGCCCAATCCCGAGGGCCGGTATACATTCCGGCCGACGCTGTGTGACTGCCTGTCCC 1074  
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320  
 DB 1075 TACAACCTACTTCTACGAGCAGTGGGATCTGCACCTCTGCTGCCCTGCACACACAA 1134  
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340  
 DB 1135 GAGGTGCACAGCAGAGATGGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1194  
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360  
 DB 1195 GTGTGCTATGGTCTGGGATGAGCAGCTTGCAGAGGTGAGGGCAGTTACCAAGTGCAT 1254  
 QY 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380  
 DB 1255 ATCCAGGAGTTTGTGGCTGCAAGAGATCTTTGGAGCCTGGCATTTCTGCCGGAGAGC 1314  
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe 400  
 DB 1315 TTTGTATGGGACCCAGCTCCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCCAAGTGT 1374  
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro 420  
 DB 1375 GAGACTCTGGAGAGATCACAGTTTACCTATACATCTCAGCATGCGCGGACAGCCTGCT 1434  
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgGlyLeuHisAsnGlyAla 440  
 DB 1435 GACCTCAGGCTCTCCAGAACCTGCAGATATCCGGGGACGAATCTGCACATGCGGCC 1494  
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
 DB 1495 TACTCGCTCACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGCTCAGTGGGAA 1554  
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480  
 DB 1555 CTGGGCACTGGAGTGGCCCTCATCCACCATACACCCACTCTGCTGCTGTCGACACAGGTG 1614  
 QY 481 ProTyrAspGlnPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
 DB 1615 CCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACTGCCACCGGCCCA 1674  
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
 DB 1675 GAGGACGAGTGTGGGGCAGGCGCTGGCTGCCACCACTGTGGCGCCGAGGCGCACTGC 1734  
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
 DB 1735 TGGGGTCCAGGGCCCAACCACTGTGTCAACTGCAGCCAGTCTCTTCCGGGCGCAGGAGTGC 1794  
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
 DB 1795 GTGGAGGAATGCCAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATCCAGGCACTGT 1854  
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580

DB 1855 TTTCCGCTGCACCTGAGTGTGAGCCGCCAGAAATGGCTGAGTACCTGTTTGGACCGGAG 1914  
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaAlaArgCys 600  
 DB 1915 GCTGACCACTGTGTGCCCTGTGCCCACTATAGGACCCCTCCCTCTGCGTGGGCCCTGTC 1974  
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
 DB 1975 CCCAGGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2034  
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640  
 DB 2035 GCGCATGTCAGCCTTGGCCCATCAATGACCACTCTCTGTGTGGACCTGGATGACAAG 2094  
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
 DB 2095 GGTGCCCCCGCAGCAGAGCCAGCCCTCTGACGTCC 2133  
 RESULT 14  
 HSERB2R  
 LOCUS Human c-erb-B-2 mRNA. 4473 bp mRNA linear PRI 30-MAR-1995  
 DEFINITION X03363  
 ACCESSION  
 VERSION  
 KEYWORDS cell surface glycoprotein; cellular oncogene; erbB-2 cellular glycoprotein; growth factor receptor; kinase; neu cellular oncogene; transmembrane protein; tyrosine kinase.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 REFERENCE 1 (bases 1 to 4473)  
 AUTHORS Yamamoto,T., Ikawa,S., Akiyama,T., Semba,K., Nomura,N., Miyajima,N., Saito,T. and Toyoshima,K.  
 TITLE Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth factor receptor  
 JOURNAL Nature 319 (6050), 230-234 (1986)  
 MEDLINE 86118663  
 PUBMED 3003577  
 REFERENCE 2  
 AUTHORS Papewalis,J., Nikitin,A.Yu. and Rajewsky,M.F.  
 TITLE G to A polymorphism at amino acid codon 655 of the human erbB-2/HER2 gene  
 JOURNAL Nucleic Acids Res. 19 (19), 5452 (1991)  
 MEDLINE 92020265  
 PUBMED 1681519  
 COMMENT The c-erb-B-2 protein shows similarity to the epidermal growth factor receptor.  
 FEATURES  
 Location/Qualifiers  
 1..4473  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /strain="cell line MKN-7"  
 /db\_xref="taxon:9606"  
 175..3942  
 /note="c-erb-B-2 precursor"  
 /codon\_start=1  
 /protein\_id="CAA27060.1"  
 /db\_xref="GI:31198"  
 /translation="MELAAALCRWGLLALLPGGAASVCTGTDMLRUPASPTHLD MURHLYQGVQVQGNLELYLPTNASLSFLDIOEVQVYLIAHNOVROPQRIRIV RGTQLFEDNYALVDNGDPLNTPPTVTCASPGGLRELRLSLTELKGVLIQRNPQ LCYQDTILMKDIFHNKNQALTLIDTNRSRACHPCSPCKGSCGSESSDCOSLRT VCAAGCARCKPLPTDCHEQCAAGCTGPKHSDCLACLFNHSIGTCLELCPALVTNT DTFESNPPEGRYTFGASCATCPYNLSTDVGSCTLVCLHNQEVATDGTGQCEKC SKPCARVCYGLGMEHLREVRVTSANIOEAGCKIFGSLAFLEPSFDGDPASNTAPL OPEQLQVETLEETGYLYISAMPDSLPDLSVFQNLQVIRPHQLINGAYSLTQIGGI SWGLRSLRHELGSGLALIHNNHILCFVHTVPMDQLFRPNHQALHTANRPECVGE LACHOLCARHSGCWGPGTQCVNCSQFLRGQCEVECRVLQGLPRETVNARHCLPCHPE COPQNGSVTCFGEADQCACAKYKDPFCVAPCSGPKPDLISYPIWKFPDEEACQ PCPINCTHSCVDLDDKCPAEQASPLTIIISAVVGILLVVVGVVFGTILIKRROOKI

```

RKYTRRLLOETELVEPLTPSGAMPNQAMRILKETELRKVKVLGSGAFGTYYKGIWI
PDGENVKIPVAIKVRENTSPKANKEILDEAYVMAGVSPVSVRLGLTCLTSTVOLVT
OLMPYGLLDHVRNKRGLSGQDLNMCQJAKMSYLEDVRLVHRDLAARNVLKSP
NHVTOEGLARLIDIDETEHYHADGGKVPKIMMALESILRRRFTHQSDVWSYGYTVRE
LMTGAKPYDGIIPAREIDPDLLEKGERLPQPPICITDVTWIMVKWMDSECRPRREL
VSESRMARDPQFVVVIONEDLGPASPLDSTFYRLSLEDDDDMDGLVDAEYLYVPOQGF
FCPPAPAGAGMVRHRSSTRSRGGDLTGLRPSSEAPRSLAPSEAGSDVFDG
DLGMAAKGLASPTHDSPLOQYSEDPVLPSETDGYVAPLTCSPQPEVYVNPQDVR
PQPSREGPLPAARPAAGATLERPKTLSPGKNGVVKDVFAGGAVENPEYLTPOGGAA
PQPHPPAFSPAFNLVYWDQDPPERGAPPSTFKGTPTAENPEYGLDVPV"
175 237
/Note="put. signal peptide (aa -21 to -1)"
238..3939
/product="put. c-erb-B-2 protein (aa 1-1234)"
376..384
/Note="pot. glycosylation site"
544..558
/Note="pot. glycosylation site"
733..741
/Note="pot. glycosylation site"
949..957
/Note="pot. glycosylation site"
1762..1770
/Note="pot. glycosylation site"
1885..1893
/Note="pot. glycosylation site"
2059..2067
/Note="pot. glycosylation site"
2353..3132
/Note="aa 727-986, seq. homologous to EGF receptor kinase domain"
2446..2454
/Note="pot. glycosylation site"
4455..4460
/Note="put. polyA signal"
4473
/Note="polyA site"
BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN

Alignment Scores:
Pred. No.: 3,45e-226 Length: 4473
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

SEQ3 (1-653) x HSERB2R (1-4473)

QY 1 MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 175 ATGGAGCTGGCGGCTTGTGGCGTGGGGGCTCTCTCTCGCCCTTGGCCCCCGGAGCC 234
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCGAGCACCACCAAGTGTGACCGGACACACATGAAGCTGCGGCTCCCTGCGAGTCCCGAG 294
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu 60
Db 295 ACCCACCCTGGACATGCTCCCGCACCTCTTACCAGGGCTGCCAGGTGGTGACAGGAAACCTG 354
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 355 GAACCTACCTACCTGCCACCAATGCCAGCTGTCCTTCTCCAGGATATCCAGAGGTG 414
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGCTACGTGCTCATCGCTCACAAACCAAGTAGGACAGGTCCTCCACTGCAGAGCTCGG 474
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 475 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTTGGCGGTGCTAGACAATGGA 534

```

```

121 AspProLeuAsnAsnThrThrProValThrGlyAlaLaserProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGCTGACAAATACACCCCTGTACAGGGGCTCCCCAGGAGGCTCGGGAGGTG 594
QY 141 GlnLeuArgSerLeuThrGluLeuLysGlyValLeuLeuIleGlnArgAsnProGln 160
Db 595 CAGCTTCAAGCCTCACAGAGATCTTAAAGAGAGGGGTCTTGATCCAGCGGAACCCAG 654
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 655 CTCTGCTACACGAGACACGATTTTGGGAAGGACATCTTCCACAGAAACACAGCTGGCT 714
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 715 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAAG 774
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 775 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTACAGAGCTGACGCGCACTGTCTGT 834
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 835 GCCGCTGGCTGTCCCGCTGCAAGGGCCACTGCCACTGACTGCTGCTGATGAGCAGGT 894
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 895 GCTGCGGCTGCACGGGCCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 955 AGTGGCATCTGTGAGCTGACTGCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 1015 TCCATGCCCCAATCCGAGGGCGGTATACATTCGCGGCCAGCTGTGTGACTGCTGCTGCTGCT 1074
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1075 TACAACCTACCTTTCTACGGAGCTGGGATCCTGACCCCTCGTCTGCGCCCTGCACAAACCA 1134
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1135 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAAGTGCAGACCCCTGTGCGCGA 1194
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1195 GTGTGCTATGCTTGGGCTGAGCAGCTTGGCAGAGGTGAGGGAGTTTACCAGTGCCTAAT 1254
QY 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1255 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGAGGCTGGCATTTCTGCCGAGAGC 1314
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1315 TTTGATGGGACCCAGCCTCCCAACACTGCCCGCTCCAGCCAGAGACGCTCCAAGTGT 1374
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420
Db 1375 GAGACTCTGGAAGAGATCACAGGTATACATCTCAGCATGCGCGGAGCAGCTCGCTGCT 1434
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1435 GACCTACGCTCTTCCAGAACCTGCAAGTATATCGGGGACCAATCTGCACAAATGCGCGC 1494
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1495 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGCTGCTGCTGCTGCTGCTGCT 1554
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1555 CTGGGAGTGGAGCTGGCCCTCATCACCATTAACACCCACTCTGCTGCTGCTGCTGCTGCTGCTGCT 1614

```

QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1615	CCCTGGGACCAAGCTTTTCGGAACCCGCCAACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1674
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1675	GAGACGAGTGTGTGGCGAGGGCTTGGCTTGCACCAAGCTGTGGCGCCGAGGGCACTGC	1734
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1735	TGGGGTCCAGGGCCACCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC	1794
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1795	GTGGAGGAATGCCGAGTACTGCAGGGCTCCCGAGGGAGTATGTGAATGCCAGGCACGT	1854
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1855	TTCGCTGGCCACCTTGAGTGTGAGCCCGAAGTGGCTCAGTGACCTGTTTGGACCGGAG	1914
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1915	GCTGACCAAGTGTGGCTGTGCCACTATAAGGACCTCCCTCTGCTGGCGCCGCTGC	1974
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu	620
Db	1975	CCCAGCGGTGAAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	2034
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2035	GGCGCATGCCAGCTTGCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAG	2094
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2095	GGCTGCCCGCCGACGACGAGCCGCTCTGACGTCC	2133
RESULT 15			
AR202597			
LOCUS	AR202597	4530 bp	DNA linear PAT 20-JUN-2002
DEFINITION	Sequence 4 from patent US 6365151.		
ACCESSION	AR202597		
VERSION	AR202597.1 GI:21498768		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 4530)		
TITLE	Halpern,M.S. and England,J.M.		
JOURNAL	Cellular immunogens comprising cognate proto-oxogenes		
FEATURES	Patent: US 6365151-A 02-APR-2002;		
source	Location/Qualifiers		
	1. 4530		
	/organism="unknown"		
BASE COUNT	922 a	1382 c	1346 g 880 t
ORIGIN			
Alignment Scores:			
Pred. No.:	3.5e-226	Length:	4530
Score:	3628.00	Matches:	653
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
SEQ3 (1-653) x AR202597 (1-4530)			
QY	1	MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla	20
Db	151	ATGGAGCTGGCGGCTTGTGCGCTGGGGGCTCTCTCGCCCTCTTGCCTCCCGGAGCC	210
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	211	GCAGGACCCCAAGTGTGCACCGGACACATGAAGCTGGCGCTCCCTGCCAGTCCCGGAG	270

QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60
Db	271	ACCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAACCTG	330
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	331	GAATCACTACCTGCCCAACCAATGCCAGCCTGTCTTCTCTCGAGGATATCCAGAGGTG	390
QY	81	GlnGlyTyrValLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	391	CAGGCTACGTGCTCATCGTCACAAAGTGAGGAGGTCCCACTGCGAGAGGTGCGG	450
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	451	ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGGTGTAGACAATGGA	510
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Db	511	GACCGCTGAACAATACCACCCCTGTACAGGGGCTCCCAAGGAGGCTTGGGGAGGTG	570
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160
Db	571	CAGCTTCAAGGCTCACAGAGATCTGAAAGAGGGGTCTTGATCCAGCGGAACCCCAAG	630
QY	161	LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAla	180
Db	631	CTCTGCTACAGGACACCATTTTGTGGAAGGACATCTTCCACAAGACAACACAGTGGCT	690
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	691	CTCACATGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAAG	750
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	751	GGTCCCCCTGCTGGGGAGAGATTCTGAGGATTGTGAGGCTTGCAGGCTGACGGGCACTG	810
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys	240
Db	811	GCCGGTGGCTGTGCCCGCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT	870
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	871	GCTCGGGCTGCACGGGCCCAAGCACCTCTGACTGCTGGCTGCCCTCCACTTCAACCCAC	930
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	931	AGTGGCATCTGTGAGCTGCATGCCCGCCCTGGTCACCTACAACACAGACACGTTTGAAG	990
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	991	TCCATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGCTCCC	1050
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	1051	TACAATACCTTTCTACGAGCTGGGATCCCTGCACCCCTGCTGCTGCCCTTGACACCAACAA	1110
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	1111	GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAGTGAGCAAGCCCTGTGCCCGA	1170
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1171	GTGTGCTATGCTTGGGCATGAGCACCTTGGAGAGGTGAGGGCAGTTTACAGTGGCCAAT	1230
QY	361	IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1231	ATCCAGGAGTTTGTGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC	1290
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1291	TTTGATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGACGACGTCCAAGTGTGTT	1350

Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420  
Db |||||||  
1351 GAGACTCTGGAGAGATACAGGTACCTATACATCTCAGCATGGCCGACAGCTGCCT 1410  
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyVala 440  
Db |||||||  
1411 GACCTCAGCGTCTTCCAGAACCTCAGAGTAATCCGGGAGCAATTCGCACAATGGCGCC 1470  
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460  
Db |||||||  
1471 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGCTGGGGCTGGCTCACTCAGGGAA 1530  
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480  
Db |||||||  
1531 CTGGGCAGTGGACTGGCCCTCATCCACATAACACCCACCTCTGCTTCGTGCACACGGTG 1590  
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500  
Db |||||||  
1591 CCCTGGACACAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1650  
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520  
Db |||||||  
1651 GAGCAGAGTGTGGCGGAGGGCTGSCCTGCCACCCAGCTGTGGCCCGGAGGCACTGC 1710  
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540  
Db |||||||  
1711 TGGGGTCCAGGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1770  
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560  
Db |||||||  
1771 GTGGAGGAATGCCAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT 1830  
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580  
Db |||||||  
1831 TTGCCGTGCCACCTGAGTGTACAGCCCAAGTGGCTCAGTACCTGTTTGGACCGGAG 1890  
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600  
Db |||||||  
1891 GCTGACCAGTGTGTGGCTGTGCCACTATAAGGACCCCTCCCTTCGCTGGCCCGCTGC 1950  
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620  
Db |||||||  
1951 CCAGCGGTGTGAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010  
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640  
Db |||||||  
2011 GGGCATGCCAGCCTTGCCCCATCACTGCACCCACTCCTCTGTGTGGACCTGGATGACAAG 2070  
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653  
Db |||||||  
2071 GGCTGCCCGCCGAGCAGAGAGCCGCTCTGACGTCC 2109

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 22:47:19 : Search time 234.794 Seconds  
(without alignments)  
3105.423 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ5  
Perfect score: 160  
Sequence: 1 ERGCPAQRASPVTFQNEGLGPASPLDSTF 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool\_p/HOLLERAN480/runat\_15102003\_131913\_20548/app\_query.fasta\_1.4685  
-DB=EST\_QFMT-fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 -CGCN\_1.12645 -runat\_15102003\_131913\_20548 -NCPU=3  
-NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST :  
1: em\_estba :  
2: em\_esthum :  
3: em\_estin :  
4: em\_estinu :  
5: em\_estov :  
6: em\_estpl :  
7: em\_estro :  
8: em\_htc :  
9: gb\_estl :  
10: gb\_est2 :  
11: gb\_htc :  
12: gb\_est3 :  
13: gb\_est4 :  
14: gb\_est5 :  
15: em\_estfun :  
16: em\_estom :  
17: em\_gss\_hum :  
18: em\_gss\_inv :  
19: em\_gss\_pln :  
20: em\_gss\_vrt :  
21: em\_gss\_fun :  
22: em\_gss\_mam :  
23: em\_gss\_mus :  
24: em\_gss\_pro :  
25: em\_gss\_fod :  
26: em\_gss\_phg :  
27: em\_gss\_vrl :  
28: gb\_gssl :

29: gb\_gss2 :\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	83	51.9	278	10	BF923905	BF923905 QV4-NT025
2	83	51.9	294	10	BF757857	BF757857 CM4-CT057
3	83	51.9	337	10	BE091682	BE091682 IL2-BT073
4	83	51.9	348	9	AI909847	AI909847 QV-BT225-
5	83	51.9	378	14	CB266492	CB266492 1005398 H
6	83	51.9	423	12	BM703962	BM703962 UI-E-CK1-
7	83	51.9	464	2	HSM067221	Bx477782 Homo sapi
8	83	51.9	496	12	BM795191	BM795191 K-EST0076
9	83	51.9	505	12	BM854421	BM854421 K-EST0136
10	83	51.9	531	12	BM830170	BM830170 K-EST0103
11	83	51.9	536	14	CB129403	CB129403 K-EST0179
12	83	51.9	547	12	BM787824	BM787824 K-EST0066
13	83	51.9	567	2	HSM077766	Bx487579 Homo sapi
14	83	51.9	574	12	BM829991	BM829991 K-EST0103
15	83	51.9	630	2	HSM073298	Bx483130 Homo sapi
16	83	51.9	691	14	CB853376	CB853376 UI-CF-FNO
17	83	51.9	767	10	BF240297	BF240297 601505830
18	83	51.9	849	14	CA489799	CA489799 AGENCOURT
19	83	51.9	852	14	CD516283	CD516283 AGENCOURT
20	83	51.9	866	14	CA488343	CA488343 AGENCOURT
21	83	51.9	893	14	CA455141	CA455141 AGENCOURT
22	83	51.9	894	14	CA454131	CA454131 AGENCOURT
23	83	51.9	899	14	CD558386	CD558386 AGENCOURT
24	83	51.9	916	13	B0845369	B0845369 AGENCOURT
25	83	51.9	919	13	B0845391	B0845391 AGENCOURT
26	83	51.9	924	14	CA489100	CA489100 AGENCOURT
27	83	51.9	932	14	CA487981	CA487981 AGENCOURT
28	83	51.9	984	14	CA489084	CA489084 AGENCOURT
29	83	51.9	1004	14	CD515356	CD515356 AGENCOURT
30	83	51.9	4715	11	AF318349	AF318349 Homo sapi
31	81	50.6	220	9	AW763060	AW763060 uc60c11.y
32	81	50.6	377	9	AW415355	AW415355 49593 MAR
33	81	50.6	483	4	BX530001	BX530001 R2PD Mus
34	81	50.6	518	10	BF039952	BF039952 Bp250015A
35	81	50.6	552	10	BF022141	BF022141 uy47c11.y
36	81	50.6	654	12	BI648646	BI648646 603277149
37	81	50.6	655	12	BI557778	BI557778 603236957
38	81	50.6	695	12	BI557170	BI557170 603239144
39	81	50.6	718	12	BI108229	BI108229 602902263
40	81	50.6	749	12	BI554781	BI554781 603236758
41	81	50.6	754	12	BI651822	BI651822 603299821
42	81	50.6	771	12	BI154609	BI154609 602904778
43	81	50.6	891	10	BF101641	BF101641 601753564
44	81	50.6	3110	11	AK031542	AK031542 Mus muscu
45	81	50.6	3372	11	BC046553	BC046553 Mus muscu

ALIGNMENTS

RESULT 1  
BF923905  
LOCUS QV4-NT0251-251100-599-all NT0251 Homo sapiens CDNA, mRNA sequence.  
DEFINITION BF923905  
ACCESSION BF923905  
VERSION BF923905.1 GI:12319793  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 278)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovskii-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

PUBLISHED

COMMENT

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV4&t2-QV4-NT0251-251100-599-all&t3-2000-11-25&t4-1)

Seq primer: puc 18 forward

High quality sequence start: 10

High quality sequence stop: 278.

Location/Qualifiers

1. .278

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="NT0251"

/notes="Organ: nervous tumor; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 52 a 78 c 79 g 69 t

ORIGIN

Alignment Scores:

Pred. No.: 0.0896 Length: 278

Score: 83.00 Matches: 16

Percent Similarity: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 2

Query Match: 51.88% Indels: 0

DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x BF923905 (1-278)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrphe 30

Db 135 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGACCTTC 188

RESULT 2

BF757857

LOCUS CM4-CT0574-101100-428-f11 CT0574 Homo sapiens cdna, mRNA sequence.

DEFINITION BF757857

VERSION BF757857.1 GI:12105757

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

1 (bases 1 to 294)

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

PUBLISHED

COMMENT

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV4&t2-QV4-NT0251-251100-599-all&t3-2000-11-25&t4-1)

Seq primer: puc 18 forward

High quality sequence start: 10

High quality sequence stop: 278.

Location/Qualifiers

1. .278

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="NT0251"

/notes="Organ: nervous tumor; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 52 a 78 c 79 g 69 t

ORIGIN

Alignment Scores:

Pred. No.: 0.0896 Length: 278

Score: 83.00 Matches: 16

Percent Similarity: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 2

Query Match: 51.88% Indels: 0

DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x BF923905 (1-278)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrphe 30

Db 135 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGACCTTC 188

RESULT 2

BF757857

LOCUS CM4-CT0574-101100-428-f11 CT0574 Homo sapiens cdna, mRNA sequence.

DEFINITION BF757857

VERSION BF757857.1 GI:12105757

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

1 (bases 1 to 294)

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

PUBLISHED

COMMENT

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV4&t2-QV4-NT0251-251100-599-all&t3-2000-11-25&t4-1)

Seq primer: puc 18 forward

High quality sequence start: 10

High quality sequence stop: 278.

Location/Qualifiers

1. .278

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="NT0251"

/notes="Organ: nervous tumor; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 52 a 78 c 79 g 69 t

ORIGIN

Alignment Scores:

Pred. No.: 0.0896 Length: 278

Score: 83.00 Matches: 16

Percent Similarity: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 2

Query Match: 51.88% Indels: 0

DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x BF923905 (1-278)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrphe 30

Db 135 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGACCTTC 188

RESULT 2

BF757857

LOCUS CM4-CT0574-101100-428-f11 CT0574 Homo sapiens cdna, mRNA sequence.

DEFINITION BF757857

VERSION BF757857.1 GI:12105757

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

1 (bases 1 to 294)

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

PUBLISHED

COMMENT

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV4&t2-QV4-NT0251-251100-599-all&t3-2000-11-25&t4-1)

Seq primer: puc 18 forward

High quality sequence start: 10

High quality sequence stop: 278.

Location/Qualifiers

1. .278

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="NT0251"

/notes="Organ: nervous tumor; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 52 a 78 c 79 g 69 t

ORIGIN

Alignment Scores:

Pred. No.: 0.0896 Length: 278

Score: 83.00 Matches: 16

Percent Similarity: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 2

Query Match: 51.88% Indels: 0

DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x BF923905 (1-278)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrphe 30

Db 135 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGACCTTC 188

RESULT 2

BF757857

LOCUS CM4-CT0574-101100-428-f11 CT0574 Homo sapiens cdna, mRNA sequence.

DEFINITION BF757857

VERSION BF757857.1 GI:12105757

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

1 (bases 1 to 294)

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

PUBLISHED

COMMENT

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV4&t2-QV4-NT0251-251100-599-all&t3-2000-11-25&t4-1)

Seq primer: puc 18 forward

High quality sequence start: 10

High quality sequence stop: 278.

Location/Qualifiers

1. .278

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="NT0251"

/notes="Organ: nervous tumor; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC



MEDLINE  
PUBMED  
COMMENT

20202663  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=st2-IL2-BT0731-240  
400-070-E05st3-2000-04-24st4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 336.

FEATURES  
source

Location/Qualifiers  
1..337  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="BT0731"  
/note="Organ: breast; Vector: puc18; Site\_1: SmaI; Site\_2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No. 196  
,716 - Ludwig Institute for Cancer Research) profiles  
into the pUC 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

BASE COUNT 72 a 101 c 96 g 68 t  
ORIGIN

Alignment Scores:

Pred. No.: 0.115 Length: 337  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x BE091682 (1-337)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 148 GTGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 95

RESULT 4  
AI909847

LOCUS QV-BT225-050599-038 BT225 Homo sapiens cDNA, mRNA linear EST 30-MAR-2000  
DEFINITION  
ACCESSION AI909847  
VERSION AI909847.1 GI:6500527  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 348)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
Brustein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare  
, M.J., Soares, F., Brentani, R.K., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
PUBMED 10737800

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/seq/gethtml.pl?t1=QVst2-QV-BT225-038.html  
st3=050599st4=1)  
Seq primer: puc 18 forward.

FEATURES  
source

Location/Qualifiers  
1..348  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/sex="female"  
/dev\_stage="Adult"  
/clone\_lib="BT225"  
/note="Organ: breast; Vector: puc18; Site\_1: SmaI; Site\_2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No. 196  
,716 - Ludwig Institute for Cancer Research) profiles  
into the pUC 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

BASE COUNT 68 a 96 c 114 g 70 t  
ORIGIN

Alignment Scores:

Pred. No.: 0.12 Length: 348  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 9 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x AI909847 (1-348)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 15 GTGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 68

RESULT 5  
CB266492

LOCUS CB266492 378 bp mRNA linear EST 20-FEB-2003  
DEFINITION 1005398 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens  
cDNA 5', mRNA sequence.

ACCESSION CB266492  
VERSION CB266492.1 GI:28441078  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 378)  
Yang, R.-Z., Shuldiner, A. and Gong, D.-W.

EST analysis of human adipose gene expression  
Unpublished  
Contact: Gong Da-Wei

Division of Endocrinology, Diabetes and Nutrition  
University of Maryland  
660 Redwood St, HH497, Baltimore, MD 21201, USA  
Tel: 410 706 1672  
Fax: 410 706 1622

Email: dgong@medicine.umaryland.edu  
PCR Primers

FORWARD: CTCGGGAAGCGCCCATTTGTTGGT

BACKWARD: AATACGACTACTATAGCGGAATTGG

Seq primer: GTTGGTACCCGGGAATTC.

FEATURES  
source

Location/Qualifiers  
1..378  
/organism="Homo sapiens"  
/mol\_type="mRNA"

```

/db_xref="taxon:9606"
/sex="Male and Female"
/tissue_type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
/notes="Vector: lambdaTriplex"
BASE COUNT      71 a 110 c 117 g  80 t
ORIGIN
Alignment Scores:
Pred. No.:      0.133      Length:      378
Score:          83.00      Matches:      16
Percent Similarity: 88.89%      Conservatives: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:      51.88%      Indels:      0
DB:              14      Gaps:      0

JUNC_SEQ8_SEQ5 (1-30) x CB266492 (1-378)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
   ||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db 124 GTGGTCATCCAGAATGAGGACTTGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 177

RESULT 6
LOCUS      BM703962
DEFINITION UI-E-CK1-afk-o-11-0-UI-r1 UI-E-CK1 Homo sapiens cDNA clone
VERSION     BM703962
KEYWORDS    EST.
SOURCE      Homo sapiens
ORGANISM    Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 423)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8989548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: Bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1..423
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CK1-afk-o-11-0-UI"
/tissue_type="Retina Foveal and Macular"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CK1"
/notes="Organ: eye; Vector: p7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-CK1 is a normalized cDNA library containing the
following tissue(s): Retina Foveal and Macular. The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated

```

to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (df)18 tail. The sequence tag for this library is GTCC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

```

BASE COUNT      78 a 129 c 131 g  85 t
ORIGIN

```

```

Alignment Scores:
Pred. No.:      0.154      Length:      423
Score:          83.00      Matches:      16
Percent Similarity: 88.89%      Conservatives: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:      51.88%      Indels:      0
DB:              12      Gaps:      0

```

JUNC\_SEQ8\_SEQ5 (1-30) x BM703962 (1-423)

```

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
   ||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db 147 GTGGTCATCCAGAATGAGGACTTGGCCCGCCAGTCCCTTGGACAGCACCTTC 200

```

RESULT 7

```

ID HSM067221 standard; RNA; EST; 464 BP.
XX
AC BX477782;
SV BX477782.1
XX
XX
DT 09-MAY-2003 (Rel. 75, Created)
DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)
XX
DE Homo sapiens mRNA; EST DKFZp686M12198_r1 (from clone DKFZp686M12198)
XX
KW EST; expressed sequence tag.
XX
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN [1]
RP 1-464
RA Bloecker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RT Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.
RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
XX
CC This is the 5' sequence of the clone insert
CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
CC Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;
CC sequenced by GBF (National Research Centre for Biotechnology
CC Ltd., Braunschweig/Germany) within the cDNA sequencing
CC consortium of the German Genome Project.
CC No sl sequence available.
CC This clone (DKFZp686M12198) is available at the RZPD in Berlin.
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,
CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX
XX
FH Key Location/Qualifiers
FH source 1..464
FT /db_xref="taxon:9606"
FT /mol_type="mRNA"
FT /organism="Homo sapiens"
FT /clone="DKFZp686M12198"
FT /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
FT DH10B; sites SfiIA + SfiIB"
FT /dev_stage="adult"

```

FT                    /tissue\_type="cdna-collection"  
 XX                    Sequence 464 BP; 87 A; 127 C; 149 G; 101 T; 0 other;

Alignment Scores:  
 Pred. No.:                    0.173                    Length:                    464  
 Score:                    83.00                    Matches:                    16  
 Percent Similarity:                    88.89%                    Conservative:                    0  
 Best Local Similarity:                    88.89%                    Mismatches:                    2  
 Query Match:                    51.88%                    Indels:                    0  
 DB:                    2                    Gaps:                    0

JUNC\_SEQ8\_SEQ5 (1-30) x HSM067221 (1-464)

OY                    13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 III  
 DB                    248 GTGGTCATCCAGATGAGGACTTGGCCCGCAGCCAGTCCCTTGGACAGCACCTTC 301

RESULT 8  
 BM795191  
 LOCUS                    K-EST0076753 S21SNU520 Homo sapiens cDNA clone S21SNU520-26-D05 5',  
 DEFINITION                    mRNA sequence.  
 ACCESSION                    BM795191  
 VERSION                    BM795191.1                    GI:19143423  
 KEYWORDS                    EST.  
 SOURCE                    Homo sapiens (human)  
 ORGANISM                    Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 AUTHORS                    Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 21C Frontier Korean EST Project 2001

TITLE                    Unpublished

JOURNAL                    Contact: Kim YS  
 COMMENT                    Genome Research Center  
                          Korea Research Institute of Bioscience & Biotechnology  
                          52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
                          Tel: +82-42-860-4470  
                          Fax: +82-42-860-4409  
                          Email: yongsung@email.kribb.re.kr  
                          Plate: 26 row: D column: 05  
                          High quality sequence stop: 496.  
 FEATURES                    Location/Qualifiers

source  
 1..496  
   /organism="Homo sapiens"  
   /mol\_type="mRNA"  
   /db\_xref="taxon:9606"  
   /clone="S21SNU520-26-D05"  
   /sex="F"  
   /tissue\_type="Stomach"  
   /cell\_type="Floating aggregates"  
   /lab\_host="SNU-520"  
   /clone\_lib="S21SNU520"  
   /organ="Organ; Stomach; Vector: pTZ18Rpl; Site:1; EcoRI:  
   Site:2; NotI; The poly (A)+ RNA was dephosphorylated with  
   bacterial alkaline phosphatase (BAP) and then decapped  
   with tabacco acid pyrophosphatase (TAP). The decapped  
   intact mRNA was ligated with DNA-RNA linker including EcoR  
   I site by treatment of T4 RNA ligase and the first strand  
   cDNA was synthesized from oligo dt-selected mRNA by  
   priming with dt-tailed vector. The dt-tailed vector was  
   adjusted to have about 60nt. The cDNA vector was  
   circularized with E. coli DNA ligase after digestion of  
   EcoRI which site is also included in vector. An RNA strand  
   converted to a DNA strand by Okayama-Berg method. The  
   obtained cDNA vectors were used for transformation of  
   competent cells E. coli Top10f<sup>+</sup> by electroporation method.  
   The cDNA libraries constructed by this method are  
   full-length enriched cDNA library."

BASE COUNT                    94 a                    134 c                    159 g                    109 t  
 ORIGIN

Alignment Scores:  
 Pred. No.:                    0.189                    Length:                    496  
 Score:                    83.00                    Matches:                    16  
 Percent Similarity:                    88.89%                    Conservative:                    0  
 Best Local Similarity:                    88.89%                    Mismatches:                    2  
 Query Match:                    51.88%                    Indels:                    0  
 DB:                    12                    Gaps:                    0

JUNC\_SEQ8\_SEQ5 (1-30) x BM795191 (1-496)

OY                    13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 III  
 DB                    334 GTGGTCATCCAGATGAGGACTTGGCCCGCAGCCAGTCCCTTGGACAGCACCTTC 387

RESULT 9  
 BM854421  
 LOCUS                    K-EST0136797 S21SNU520 Homo sapiens cDNA clone S21SNU520-57-G11 5',  
 DEFINITION                    mRNA sequence.  
 ACCESSION                    BM854421  
 VERSION                    BM854421.1                    GI:19210820  
 KEYWORDS                    EST.  
 SOURCE                    Homo sapiens (human)  
 ORGANISM                    Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 AUTHORS                    Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 21C Frontier Korean EST Project 2001

TITLE                    Unpublished

JOURNAL                    Contact: Kim YS  
 COMMENT                    Genome Research Center  
                          Korea Research Institute of Bioscience & Biotechnology  
                          52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
                          Tel: +82-42-860-4470  
                          Fax: +82-42-860-4409  
                          Email: yongsung@email.kribb.re.kr  
                          Plate: 57 row: G column: 11  
                          High quality sequence stop: 505.  
 FEATURES                    Location/Qualifiers

source  
 1..505  
   /organism="Homo sapiens"  
   /mol\_type="mRNA"  
   /db\_xref="taxon:9606"  
   /clone="S21SNU520-57-G11"  
   /sex="F"  
   /tissue\_type="Stomach"  
   /cell\_type="Floating aggregates"  
   /lab\_host="SNU-520"  
   /clone\_lib="S21SNU520"  
   /organ="Organ; Stomach; Vector: pTZ18Rpl; Site:1; EcoRI:  
   Site:2; NotI; The poly (A)+ RNA was dephosphorylated with  
   bacterial alkaline phosphatase (BAP) and then decapped  
   with tabacco acid pyrophosphatase (TAP). The decapped  
   intact mRNA was ligated with DNA-RNA linker including EcoR  
   I site by treatment of T4 RNA ligase and the first strand  
   cDNA was synthesized from oligo dt-selected mRNA by  
   priming with dt-tailed vector. The dt-tailed vector was  
   adjusted to have about 60nt. The cDNA vector was  
   circularized with E. coli DNA ligase after digestion of  
   EcoRI which site is also included in vector. An RNA strand  
   converted to a DNA strand by Okayama-Berg method. The  
   obtained cDNA vectors were used for transformation of  
   competent cells E. coli Top10f<sup>+</sup> by electroporation method.  
   The cDNA libraries constructed by this method are  
   full-length enriched cDNA library."

BASE COUNT                    95 a                    138 c                    161 g                    111 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.193 Length: 505  
 Score: 83.00 Matches: 16  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 51.88% Indels: 0  
 DB: 12 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x BM854421 (1-505)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 Db 334 GTGGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 387  
 RESULT 10  
 BM830170  
 LOCUS K-EST0103592 S21SNU520s1 531 bp mRNA linear EST 06-MAR-2002  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BM830170  
 VERSION BM830170.1 GI:19186579  
 KEYWORDS Est.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 531)  
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@email.kribb.re.kr  
 Plate: 19 row: D column: 03  
 High quality sequence stop: 531.  
 Location/Qualifiers

## FEATURES

source

1. 531  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S21SNU520s1-19-D03"  
 /sex="F"  
 /tissue\_type="Stomach"  
 /cell\_type="floating aggregates"  
 /cell\_line="SNU-520"  
 /lab\_host="Top10f"  
 /clone\_lib="S21SNU520s1"  
 /note="Organ: Stomach; Vector: pTZ18RP1; Site.1: EcoRI;  
 Site.2: NotI; The poly (A)+ RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then decapped  
 with tobacco acid pyrophosphatase (TAP). The decapped  
 intact mRNA was ligated with DNA-RNA linker including EcoR  
 I site by treatment of T4 RNA ligase and the first strand  
 cDNA was synthesized from oligo dt-selected mRNA by  
 priming with dt-tailed vector. The dt-tailed vector was  
 adjusted to have about 60nt. The cDNA vector was  
 circularized with E. coli DNA ligase after digestion of  
 EcoRI which site is also included in vector. An RNA strand  
 converted to a DNA strand by Okayama-Berg method. The  
 obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10f' by electroporation method.  
 The cDNA libraries constructed by this method are  
 full-length enriched cDNA library. After analyzing and  
 sequencing about 2,000 ~ 3,000 colonies in original cDNA  
 library, the abundant cDNAs were selected and amplified by

PCR reaction using vector region primer including T7  
 promotor as 5' primer and N(dt)14 as 3' primer. The PCR  
 products were used as template for synthesis of  
 biotinylated single stranded RNA by in vitro transcription  
 reaction. The synthesized RNA probes were hybridized with  
 antisense single stranded cDNAs prepared from original  
 library and incubated with avidin-gel. After removing  
 DNA-RNA hybrids by centrifuge, the subtracted cDNA  
 libraries were constructed by transformation of the  
 remaining DNA into competent cells E. coli Top10f' with  
 electroporation method."

BASE COUNT 102 a 149 c 166 g 114 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.206 Length: 531  
 Score: 83.00 Matches: 16  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 51.88% Indels: 0  
 DB: 12 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x BM830170 (1-531)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 Db 334 GTGGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 387

## RESULT 11

CB129403

LOCUS

DEFINITION

CB129403

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

21C Frontier Korean EST Project 2001  
 Unpublished  
 Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@email.kribb.re.kr  
 Plate: 5 row: C column: 08  
 High quality sequence stop: 536.  
 Location/Qualifiers

## FEATURES

source

1. 536  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="C1SNU17-5-C08"  
 /sex="F"  
 /tissue\_type="uterine"  
 /cell\_type="epithelial"  
 /cell\_line="SNU-17"  
 /lab\_host="Top10f"  
 /clone\_lib="C1SNU17"  
 /note="Organ: Cervix; Vector: pCNS-D2; Site.1: EcoRI;  
 Site.2: NotI; The poly (A)+ RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then decapped  
 with tobacco acid pyrophosphatase (TAP). The decapped  
 intact mRNA was ligated with DNA-RNA linker including  
 EcoRI site by treatment of T4 RNA ligase and the first  
 strand cDNA was synthesized from oligo dt-selected mRNA by  
 priming with dt-tailed vector. The dt-tailed vector was  
 adjusted to have about 60nt. The cDNA vector was  
 circularized with E. coli DNA ligase after digestion of  
 EcoRI which site is also included in vector. An RNA strand  
 converted to a DNA strand by Okayama-Berg method. The  
 obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10f' by electroporation method.  
 The cDNA libraries constructed by this method are  
 full-length enriched cDNA library. After analyzing and  
 sequencing about 2,000 ~ 3,000 colonies in original cDNA  
 library, the abundant cDNAs were selected and amplified by

strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F<sup>+</sup> by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 105 a 150 c 167 g 114 t

ORIGIN

Alignment Scores: Pred. No.: 0.209 Length: 536  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 14 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x CB129403 (1-536)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 339 GTGGTCATCCAGAAATGAGGACTTGGCCCGCCAGCAGCTCCCTGGACAGCACCTTC 392

RESULT 12

BM787824

LOCUS

DEFINITION K-EST0066898 SL1SN01 Homo sapiens cDNA clone S11SN01-23-G01 5',  
547 bp mRNA linear EST 05-MAR-2002

BM787824

ACCESSION BM787824.1 GI:19136056

VERSION

KEYWORDS

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE

AUTHORS

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 23 row: G column: 01

High quality sequence stop: 547.

Location/Qualifiers

1. 547

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S11SN01-23-G01"

/sex="M"

/tissue\_type="Stomach"

/cell\_type="Lymphoblast-like"

/lab\_host="Top10F"

/clone\_lib="S11SN01"

/note="Organ: Stomach; Vector: pME18-FL3; Site.1: XhoI;

Site.2: XhoI; The poly (A)+ RNA was dephosphorylated with

bacterial alkaline phosphatase (BAP) and then decapped

with tobacco acid pyrophosphatase (TAP). The decapped

intact mRNA was ligated with DNA-RNA linker including SfiI

site by treatment of T4 RNA ligase and the first strand

cDNA was synthesized with Superscript II using SfiI

oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into DraIII- digested pME18s-FL3 vector. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F<sup>+</sup> by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 102 a 187 c 163 g 95 t

ORIGIN

Alignment Scores: Pred. No.: 0.214 Length: 547  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x BM787824 (1-547)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 39 GTGGTCATCCAGAAATGAGGACTTGGCCCGCCAGCAGCTCCCTGGACAGCACCTTC 92

RESULT 13

HSM077766

ID HSM077766 standard; RNA; EST; 567 BP.

XX AC BX487579;

XX SV BX487579.1

XX DT 09-MAY-2003 (Rel. 75, Created)

XX DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)

XX Homo sapiens mRNA; EST DKFZp686M16257\_r1 (from clone DKFZp686M16257)

XX EST; expressed sequence tag.

XX Homo sapiens (human)

XX Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia;

XX Eutheria; Primates; Catarrhini; Hominidae; Homo.

XX [1]

RP 1-567

RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,

Han M., Wiemann S.;

Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.

MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY.

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by BMFZ (Biomedical Research Center at the

Heinrich-Heine-University, Duesseldorf/Germany) within the cDNA

sequencing consortium of the German Genome Project.

No sl sequence available.

This clone (DKFZp686M16257) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,

14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

XX Key Location/Qualifiers

1. 567

/db\_xref="taxon:9606"

/mol\_type="mRNA"

/organism="Homo sapiens"

/clone="DKFZp686M16257"

/clone\_lib="686 (synonym: hlcc3). Vector pSport1\_sfi; host

DH10B; sites SfiIA + SfiIB"

/dev\_stage="adult"



FT /clone="DKF2p686N09234"  
FT /clone\_lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host  
FT DH10B; sites SfiIA + SfiIB"  
FT /dev\_stage="adult"  
FT /tissue\_type="cDNA-collection"  
XX  
SQ Sequence 630 BP; 124 A; 174 C; 196 G; 133 T; 3 other;

Alignment Scores:  
Pred. No.: 0.257 Length: 630  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 2 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x HSM073298 (1-630)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
||| ||||||||||||||||||||||||||||||||||||||||||||  
Db 439 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 492

Search completed: October 16, 2003, 17:04:07  
Job time : 235.794 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 16, 2003, 11:08:07 ; Search time 31.0111 Seconds  
(without alignments)  
2540.503 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ5

Perfect score: 160

Sequence: 1 ERGCCPAERASPTVTFQNEGLGASPILDSTF 30

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO.spool.p/HOLLERANA480/runat\_15102003\_131915\_20662/app\_query.fasta\_1.4685  
-DB=Published\_Applications\_NA -QPMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USPR=HOLLERANA480 @CGN.1.1.397 @runat\_15102003\_131915\_20662  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published\_Applications\_NA:  
1: /cgn2.6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2.6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2.6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2.6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2.6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2.6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2.6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2.6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2.6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2.6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2.6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2.6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2.6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2.6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2.6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
16: /cgn2.6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
17: /cgn2.6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	ID	Description
1	83	51.9	1115	14	US-10-102-806-165 Sequence 165, App

2	83	51.9	1713	12	US-10-378-393-14 Sequence 14, Appl
3	83	51.9	1755	10	US-09-930-125-6 Sequence 6, Appl
4	83	51.9	1767	10	US-09-930-125-4 Sequence 4, Appl
5	83	51.9	1773	10	US-09-930-125-7 Sequence 7, Appl
6	83	51.9	1806	10	US-09-930-125-5 Sequence 5, Appl
7	83	51.9	2411	12	US-10-378-393-10 Sequence 10, Appl
8	83	51.9	3765	12	US-10-207-498-5 Sequence 5, Appl
9	83	51.9	3768	9	US-09-811-123-8 Sequence 8, Appl
10	83	51.9	3768	9	US-09-811-115-2 Sequence 9, Appl
11	83	51.9	3768	10	US-09-854-356-9 Sequence 1, Appl
12	83	51.9	3768	10	US-09-930-125-1 Sequence 1, Appl
13	83	51.9	3768	12	US-10-313-644-1 Sequence 1, Appl
14	83	51.9	4473	11	US-09-441-411-5 Sequence 5, Appl
15	83	51.9	4473	12	US-10-101-510-81 Sequence 81, Appl
16	83	51.9	4473	14	US-10-146-473-32 Sequence 32, Appl
17	83	51.9	4473	14	US-10-207-655-44 Sequence 44, Appl
18	83	51.9	4530	10	US-09-877-177-11 Sequence 11, Appl
19	83	51.9	4530	12	US-10-007-926A-119 Sequence 119, App
20	83	51.9	4530	12	US-10-101-510-124 Sequence 124, App
21	83	51.9	4530	12	US-10-338-730-1 Sequence 1, Appl
22	83	51.9	4530	14	US-10-177-293-125 Sequence 125, App
23	83	51.9	4543	10	US-09-769-508-1 Sequence 1, Appl
24	83	51.9	4606	12	US-09-971-392-70 Sequence 70, Appl
25	83	51.9	4642	14	US-10-198-846-10896 Sequence 10896, A
26	83	51.9	9274	9	US-09-811-123-7 Sequence 7, Appl
27	83	51.9	9274	9	US-09-811-115-1 Sequence 1, Appl
28	81	50.6	3771	10	US-09-854-356-11 Sequence 11, Appl
29	81	50.6	3955	10	US-09-870-759-117 Sequence 117, App
30	81	50.6	3955	10	US-09-854-356-10 Sequence 10, Appl
31	81	50.6	3955	12	US-09-751-708A-117 Sequence 117, App
32	66	41.2	201	13	US-10-109-213-3 Sequence 3, Appl
33	58.5	36.6	143601	12	US-09-855-824-3 Sequence 25, Appl
34	58	36.2	1938	12	US-10-422-264-25 Sequence 27, Appl
35	58	36.2	2316	12	US-10-422-264-27 Sequence 23, Appl
36	58	36.2	2604	12	US-10-422-264-23 Sequence 1, Appl
37	58	36.2	3320	10	US-09-862-658-1 Sequence 29, Appl
38	58	36.2	3320	14	US-10-175-696-22 Sequence 22, Appl
39	58	36.2	3384	12	US-10-422-264-29 Sequence 29, Appl
40	57.5	35.9	363	11	US-09-918-995-28706 Sequence 28706, A
41	57	35.6	23432	9	US-09-764-869-1332 Sequence 1332, Ap
42	57	35.6	23432	14	US-10-091-504-1332 Sequence 1332, Ap
43	56.5	35.3	1557	10	US-09-774-434-6 Sequence 6, Appl
44	56.5	35.3	22452	10	US-09-764-868-1487 Sequence 1487, Ap
45	56.5	35.3	22452	10	US-09-764-868-1489 Sequence 1489, Ap

ALIGNMENTS

RESULT 1  
US-10-102-806-165  
; Sequence 165, Application US/10102806  
; Publication No. US20030054421A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA103PIC1  
; CURRENT APPLICATION NUMBER: US/10/102,806  
; CURRENT FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 09/925,298  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05881  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 846  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 165  
; LENGTH: 1115  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (390)



```
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-7

Alignment Scores:
Pred. No.: 0.00233 Length: 1773
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-930-125-7 (1-1773)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 961 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 1014

RESULT 6
US-09-930-125-5
; Sequence 5, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1806
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-5

Alignment Scores:
Pred. No.: 0.00238 Length: 1806
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-930-125-5 (1-1806)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 994 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 1047

RESULT 7
US-10-378-393-10
; Sequence 10, Application US/10378393
; Publication No. US20030182668A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVAT
; FILE REFERENCE: D0254 NP
; CURRENT APPLICATION NUMBER: US/10/378,393
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,889
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-378-393-10

Alignment Scores:
Pred. No.: 0.00327 Length: 2411
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-10-378-393-10 (1-2411)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 1602 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 1655

RESULT 8
US-10-207-498-5
; Sequence 5, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HERGULIN AND HER3
; FILE REFERENCE: 30448.103-US-01
; CURRENT APPLICATION NUMBER: US/10/207,498
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-207-498-5

Alignment Scores:
Pred. No.: 0.00532 Length: 3765
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-10-207-498-5 (1-3765)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
```

```

Db      2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 3012
      ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
RESULT 9
US-09-811-123-8
; Sequence 8, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
; FILE REFERENCE: ANTIBODY-MAYTANSINOID CONJUGATES
; CURRENT APPLICATION NUMBER: 09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-123-8

Alignment Scores:
Pred. No.: 0.00532 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 9 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-811-123-8 (1-3768)

Qy      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
      ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 3012
      ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
RESULT 10
US-09-811-115-2
; Sequence 2, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-115-2

Alignment Scores:
Pred. No.: 0.00532 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 9 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-811-115-2 (1-3768)

```

```

Qy      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
      ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 3012
      ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
RESULT 11
US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc.feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc.feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc.feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc.feature
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

Alignment Scores:
Pred. No.: 0.00532 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-854-356-9 (1-3768)

Qy      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
      ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 3012
      ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
RESULT 12
US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.

```

```

: APPLICANT: Scholler, Nathalie B.
: APPLICANT: Disis, Mary L.
: APPLICANT: Hellstrom, Ingegerd
: APPLICANT: Hellstrom, Karl Erik
: TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
: FILE REFERENCE: 730033.409
: CURRENT APPLICATION NUMBER: US/09/441.411
: CURRENT FILING DATE: 1999-11-16
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 4473
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-441-411-5

Alignment Scores:
Pred. No.: 0.00642 Length: 4473
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 11 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-441-411-5 (1-4473)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3133 GTGTCATCATCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 3186

RESULT 15
US-10-101-510-81
: Sequence 81, Application US/10101510
: Publication No. US20030148295A1
: GENERAL INFORMATION:
: APPLICANT: WAN, JACKSON
: APPLICANT: WANG, YIXIN
: TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
: FILE REFERENCE: 15117.0012
: CURRENT APPLICATION NUMBER: US/10/101.510
: CURRENT FILING DATE: 2002-03-20
: PRIOR APPLICATION NUMBER: 60/276,947
: PRIOR FILING DATE: 2001-03-20
: NUMBER OF SEQ ID NOS: 805
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 81
: LENGTH: 4473
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-101-510-81

Alignment Scores:
Pred. No.: 0.00642 Length: 4473
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-10-101-510-81 (1-4473)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3133 GTGTCATCATCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 3186

Search completed: October 17, 2003, 03:53:14
Job time : 34.0111 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 22:55:19 ; Search time 7.60488 Seconds

(without alignments)  
1741.185 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ5

Perfect score: 150

Sequence: 1 ERGCPAQRASPVTFQNDLGASPLDSTF 30

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/HOLLERAN480/runat\_15102003\_131913\_20570/app\_query.fasta\_1.4685  
-DB=Issued\_Patents\_NA -QPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosu62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFTW=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480.@CGL\_1\_1403.@runat\_15102003\_131913\_20570 -NCPU=6 -ICPU=3  
-NO\_MMAPP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	51.9	3768	2	US-08-625-101-1
2	83	51.9	3768	2	US-08-356-786-1
3	83	51.9	4473	2	US-09-048-804-1
4	83	51.9	4473	3	US-09-056-105-26
5	83	51.9	4530	1	US-08-229-515A-9
6	83	51.9	4530	1	US-08-645-865-9
7	83	51.9	4530	4	US-09-167-322-4
8	83	51.9	4530	4	US-09-527-487-1
9	83	51.9	4530	4	US-09-877-177A-11
10	81	50.6	3955	1	US-08-229-515A-14
11	81	50.6	3955	1	US-08-645-865-14
12	68	42.5	2385	2	US-09-146-283-3

13	68	42.5	2385	3	US-08-579-823A-3	Sequence 3, Appli
14	68	42.5	2385	3	US-09-344-195-3	Sequence 3, Appli
15	66	41.2	153	3	US-08-776-251-3	Sequence 3, Appli
16	66	41.2	201	4	US-09-200-355-3	Sequence 3, Appli
17	66	41.2	816	3	US-08-776-251-10	Sequence 10, Appl
18	58	36.2	1938	4	US-09-547-435-25	Sequence 25, Appl
19	58	36.2	2316	4	US-09-547-435-27	Sequence 27, Appl
20	58	36.2	2604	4	US-09-547-435-23	Sequence 23, Appl
21	58	36.2	3384	4	US-09-547-435-29	Sequence 29, Appl
22	56.5	35.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
23	56.5	35.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
24	56	35.0	1368	4	US-09-328-352-430	Sequence 430, App
25	55.5	34.7	2691	4	US-09-020-743-1	Sequence 1, Appli
26	54.5	34.1	536165	4	US-09-214-808-1	Sequence 1, Appli
27	54	33.8	42571	4	US-09-810-347-3	Sequence 3, Appli
28	53.5	33.4	634	1	US-08-451-947-1	Sequence 1, Appli
29	53.5	33.4	634	2	US-08-424-826A-1	Sequence 1, Appli
30	53.5	33.4	634	3	US-08-928-694-1	Sequence 1, Appli
31	53.5	33.4	634	4	US-08-450-842-1	Sequence 1, Appli
32	53.5	33.4	634	4	US-08-451-390-1	Sequence 1, Appli
33	53.5	33.4	634	5	PCT-US91-06950-1	Sequence 1, Appli
34	53.5	33.4	864	3	US-09-022-983-3	Sequence 3, Appli
35	53.5	33.4	1404	1	US-07-796-106-22	Sequence 22, Appl
36	53.5	33.4	2340	3	US-09-022-983-4	Sequence 4, Appli
37	53.5	33.4	2477	3	US-09-490-692-3	Sequence 3, Appli
38	53	33.1	960	3	US-08-651-136C-1	Sequence 1, Appli
39	53	33.1	960	4	US-09-229-911A-1	Sequence 1, Appli
40	53	33.1	5046	4	US-09-548-938A-5	Sequence 5, Appli
41	53	33.1	111282	4	US-09-754-250-3	Sequence 3, Appli
42	51.5	32.2	3225	4	US-09-358-856C-2	Sequence 2, Appli
43	51	31.9	4092	3	US-09-306-595C-5	Sequence 5, Appli
44	51	31.9	4092	4	US-09-925-388-5	Sequence 5, Appli
45	51	31.9	7672	4	US-09-220-132-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-08-625-101-1

; Sequence 1, Application US/08625101

; Patent No. 5869445

; GENERAL INFORMATION:

; APPLICANT: Cheever, Martin A.

; APPLICANT: Disis, Mary L.

; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE

; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION

; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/625,101

; FILING DATE: 01-APR-1996

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Sharkey, Richard G.

; REGISTRATION NUMBER: 32,629

; REFERENCE/DOCKET NUMBER: 920010.448C7

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 1:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3765
US-08-625-101-1

Alignment Scores:
Pred. No.: 0.00745 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-08-625-101-1 (1-3769)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATCGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 3012

RESULT 2
US-08-356-786-1
; Sequence 1, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3768
; OTHER INFORMATION: /note= "product = "cerB-b2""
```

```
US-08-356-786-1

Alignment Scores:
Pred. No.: 0.00745 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-08-356-786-1 (1-3768)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATCGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 3012

RESULT 3
US-09-048-804-1
; Sequence 1, Application US/09048804
; Patent No. 5968748
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
; TITLE OF INVENTION: HUMAN HER-2 EXPRESSION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748ris LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/048,804
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legaard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2913
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4473 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; ANTI-SENSE: NO
US-09-048-804-1

Alignment Scores:
Pred. No.: 0.0092 Length: 4473
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-048-804-1 (1-4473)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3133 GTGGTCATCCAGATCGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 3186

RESULT 4
US-09-056-105-26
```



; Sequence 26, Application US/09056105  
; Patent No. 6287569  
; GENERAL INFORMATION:  
; APPLICANT: KIPPS, THOMAS J.  
; APPLICANT: WU, YUNQI  
; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR  
; FILE REFERENCE: 233/221  
; CURRENT APPLICATION NUMBER: US/09/056,105  
; EARLIER FILING DATE: 1998-04-06  
; EARLIER APPLICATION NUMBER: 50/043,467  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 26  
; LENGTH: 4473  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-056-105-26

Alignment Scores:  
Pred. No.: 0.0092 Length: 4473  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 3 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x US-09-056-105-26 (1-4473)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 3133 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCTCCCTTGGACAGCACCTTC 3186

RESULT 5  
US-08-229-515A-9  
; Sequence 9, Application US/08229515A  
; Patent No. 5518885  
; GENERAL INFORMATION:  
; APPLICANT: RAZIYUDDIN  
; APPLICANT: SARKAR, FAZLUL H  
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN  
; TITLE OF INVENTION: NEOPLASTIC DISEASE  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG PC  
; STREET: 127 Peachtree Street, Suite 1200  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: usa  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/229,515A  
; FILING DATE: 19 APR 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PERRYMAN, DAVID G  
; REGISTRATION NUMBER: 33,438  
; REFERENCE/DOCKET NUMBER: 1414.608  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404-688-0770  
; TELEFAX: 404-688-9880  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4530 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-229-515A-9

Alignment Scores:  
Pred. No.: 0.00935 Length: 4530  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 1 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x US-08-229-515A-9 (1-4530)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCTCCCTTGGACAGCACCTTC 3162

RESULT 6  
US-08-645-865-9  
; Sequence 9, Application US/08645865  
; Patent No. 5654406  
; GENERAL INFORMATION:  
; APPLICANT: RAZIYUDDIN  
; APPLICANT: SARKAR, FAZLUL H  
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN  
; TITLE OF INVENTION: NEOPLASTIC DISEASE  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG PC  
; STREET: 127 Peachtree Street, Suite 1200  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: usa  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/645,865  
; FILING DATE: 14 MAY 1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PERRYMAN, DAVID G  
; REGISTRATION NUMBER: 33,438  
; REFERENCE/DOCKET NUMBER: 1414.608  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404-688-0770  
; TELEFAX: 404-688-9880  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4530 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-645-865-9

Alignment Scores:  
Pred. No.: 0.00935 Length: 4530  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 1 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x US-08-645-865-9 (1-4530)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCTCCCTTGGACAGCACCTTC 3162

RESULT 7  
US-09-167-322-4

```
; Sequence 4, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-167-322-4

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-167-322-4 (1-4530)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 8
US-09-527-487-1
; Sequence 1, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(3915)
US-09-527-487-1

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-527-487-1 (1-4530)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 9
US-09-877-177A-11
; Sequence 11, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-11

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x US-09-877-177A-11 (1-4530)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 10
US-08-229-515A-14
; Sequence 14, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

; TOPOLOGY: linear
US-08-645-865-14

Alignment Scores:
Pred. No.:      0.0158      Length:      3955
Score:          81.00       Matches:     15
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:    50.62%       Indels:      0
DB:             1           Gaps:        0

JUNC_SEQ8_SEQ5 (1-30) x US-08-645-865-14 (1-3955)
QY      1 GluArgGlyCysProAlaGluGlnArGalaSerProValThrPhe 15
Db      1946 GAACGAGGCTGCCCGACGACGAGCCAGCCGGTGCATTC 1990
|||||
RESULT 12
US-09-146-283-3
; Sequence 3, Application US/09146283
; Patent No. 5976546
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-09-146-283-3

Alignment Scores:
Pred. No.:      0.766      Length:      2385
Score:          68.00       Matches:     17
Percent Similarity: 66.67%   Conservative: 3
Best Local Similarity: 56.67% Mismatches:     6
Query Match:     42.50%       Indels:      4
DB:             2           Gaps:        2

JUNC_SEQ8_SEQ5 (1-30) x US-09-146-283-3 (1-2385)

```

Qy	1	GluArgGlyCysProAlaGluInrArgAlaSerProValThrPheGlnAsnGluAspLeu	20
Db	1925	GACAGGGCTGCCCGCCGACAGAGAGAGACCCTCTGCAGTCCCTCGAG-----	1975
Qy	21	GlyProAla---SerProLeuAspSerThr	29
Db	1976	GCACCGCGCCGCTCGGCCAGCCCCAGCACCA	2005

```

, RESULT 13
, US-08-579-823A-3
, ; Sequence 3, Application US/08579823A
, ; Patent No. 6080409
, ; GENERAL INFORMATION:
, ; APPLICANT: Laus, Reiner
, ; APPLICANT: Ruegg, Curtis L.
, ; APPLICANT: Wu, Hongyu
, ; TITLE OF INVENTION: Immunostimulatory Composition and Method
, ; NUMBER OF SEQUENCES: 10
, ; CORRESPONDENCE ADDRESS:
, ; ADDRESSEE: Dehlinger & Associates
, ; STREET: 350 Cambridge Ave. Suite 250
, ; CITY: Palo Alto
, ; STATE: CA
, ; COUNTRY: USA

```

217. 34300  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/579,823A  
APPLICATION NUMBER: 03-DEC-1998  
FILING DATE: 536  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Judge, Linda R.  
REGISTRATION NUMBER: 42,702  
REFERENCE/DOCKET NUMBER: 7636-0010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-324-0880  
TELEFAX: 650-324-0960  
INFORMATION FOR SEQ ID NO: 3:

[illegible]

Alignment Scores:	
Pred. No.:	0.766
Score:	68.00
Percent Similarity:	66.67%
Best Local Similarity:	56.67%
Query Match:	42.50%
DB:	3
Gaps:	2
Indels:	4
Mismatches:	6
Conservative:	3
Matches:	17
Length:	2385

JUNC\_SEQ8\_SEQ5 (1-30) x US-08-579-823A-3 (1-2385)

Qy	1	GLuArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu	20
Db	1925	GACAGGGGTGCCCCCGGACGACAGAGCAGCCCTCTGACGTCCTCGAG-----	1975
Qy	21	GlyProAla---SerProLeuAspSerThr	29
Db	1976	GCACCGCGCCGGCTCGCCAGCCCGACGAC	2005

RESULT 14  
US-09-344-195-3  
; Sequence 3, Application US/09344195  
; Patent No. 6210662  
; GENERAL INFORMATION:  
; APPLICANT: Laus, Reiner  
; Ruegg, Curtis L.  
; Wu, Hongyu  
; TITLE OF INVENTION: Immunostimulatory Compositions  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Ave. Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/344,195  
; FILING DATE: 24-Jun-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/146,283  
; FILING DATE: 03-SEPT-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Judge, Linda R.  
; REGISTRATION NUMBER: 42,702  
; REFERENCE/DOCKET NUMBER: 7636-0010.21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-0880  
; TELEFAX: 650-324-0960  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2385 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: homo sapiens  
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene;  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-344-195-3

Alignment Scores:	
Pred. NO.:	0.766
Score:	68.00
Percent Similarity:	66.67%
Best Local Similarity:	56.67%
Query Match:	42.50%
DB:	3
Length:	2385
Matches:	17
Conservative:	3
Mismatches:	6
Indels:	2
Gaps:	2

JUNC\_SEQ8\_SEQ5 (1-30) x US-09-344-195-3 (1-2385)

Qy		1	GluArgGlyCysProAlaGluInrA9AlaSerProValThrPheGlnnAsnGluAaspLeu	20
	:	:	:	:
Db		1925	GACAAGGGCTGCCCGCCGACAGACAGCCCCCTGTGACGTCCCTCGAG-----	1975
Qy		21	GlyProAla---SerProLeuAaspSerThr	29
	:	:	:	:
Db		1976	GCACCCGCCCGCTCGCCAGCCCAGCAC	2005

RESULT 15  
US-08-776-251-3  
; Sequence 3, Application US/08776251

Search completed: October 16, 2003, 17:16:45  
Job time : 16.6049 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 28.4941 Seconds  
(without alignments)  
2842.104 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ5

Perfect score: 160  
Sequence: 1 ERGCPAEQASPVTFQNEGLGPASPLDSTF 30

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-framer-p2n.model -DEV-xlp  
-Q/cn2\_1/USPTO\_spo1\_p/HOLLERAN480/runat\_15102003\_131912\_20526/app\_query\_fasta\_1.4695  
-DB=N\_Geneseq\_19Jun03 -OFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPLC=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 -CGCN 1.1 1758 -runat\_15102003\_131912\_20526 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_19Jun03.\*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	147	91.9	2763	24	ABA92252	Mouse Her-2/neu ex
2	147	91.9	2781	24	ABA92253	Her-2/neu extracel
3	83	51.9	1115	21	AAF21778	Human breast and o
4	83	51.9	1755	24	AAD32746	Human cDNA for the
5	83	51.9	1767	24	AAD32744	Human cDNA for the
6	83	51.9	1773	24	AAD32747	Human cDNA for the
7	83	51.9	1806	24	AAD32745	Human cDNA for the
8	83	51.9	3600	21	AA89736	Human HER-2/neu co
9	83	51.9	3578	24	ABK86207	cDNA encoding huma
10	83	51.9	3768	17	AAT40739	HER-2/neu oncogene
11	83	51.9	3768	20	AA01912	Human HER-2/neu on
12	83	51.9	3768	21	AA09455	Human heregulin 2
13	83	51.9	3768	22	AAH23392	Human HER-2/neu pr
14	83	51.9	3768	24	ABZ35744	Human ERBB2 polynu
15	83	51.9	3768	24	ABX0987	Human ERBB2 DNA fr
16	83	51.9	3768	24	AA043935	Human HER-2 cDNA.
17	83	51.9	3768	24	AA043986	Human ERBB2 antigen
18	83	51.9	3768	24	ABV78168	Human ERBB2 DNA SE
19	83	51.9	3768	24	AA032743	Human Her-2/neu pr
20	83	51.9	3768	24	ABA92250	Human Her-2/neu CD
21	83	51.9	3768	24	ABK10730	Human Her-2/neu DN
22	83	51.9	3768	24	ABL91709	Human polynucleoti
23	83	51.9	3768	24	ABK14058	Human HER2 (Erbb2)
24	83	51.9	4299	14	AA046083	Sequence encoding the
25	83	51.9	4472	21	AA14812	cDNA encoding the
26	83	51.9	4473	19	ABQ76220	Human tumour anti
27	83	51.9	4473	20	AAZ31071	HER-2 nucleic acid
28	83	51.9	4473	24	ABZ34969	Human gene express
29	83	51.9	4473	24	AA038904	Human Her-2 DNA.
30	83	51.9	4530	16	AAT01585	Her-2/neu (ERBB2/c
31	83	51.9	4530	18	AAT71253	Human ER2 gene.
32	83	51.9	4530	21	AAZ60815	Nucleotide sequenc
33	83	51.9	4530	22	AAD19731	Human tyrosine kin
34	83	51.9	4530	24	ABV35012	Human gene express
35	83	51.9	4530	24	ABV94128	Breast carcinoma r
36	83	51.9	4530	24	ABN85585	Human HER2-neu SQ
37	83	51.9	4530	24	ABK83918	Human cDNA differe
38	83	51.9	4530	25	ACC50139	Breast cancer asso
39	83	51.9	4530	25	ABQ83856	Human Her2/Neu enc
40	83	51.9	9274	22	AAF24297	HER2 transgene pla
41	83	51.9	9274	24	AAD43934	HER-2 transgene pl
42	83	51.9	9274	24	ABK14057	Human HER2 (Erbb2)
43	81	50.6	3771	21	AA89737	Mouse Her-2/neu CD
44	81	50.6	3771	22	AAH42210	Nucleotide sequenc
45	81	50.6	3771	24	ABA92251	Mouse Her-2/neu CD

ALIGNMENTS

RESULT 1  
ABA92252  
ID ABA92252 standard; cDNA; 2763 BP.  
XX  
AC ABA92252;  
XX  
DT 17-JUN-2002 (first entry)  
XX  
DE Mouse Her-2/neu extracellular-phosphorylation domain fusion cDNA.  
XX  
KW Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;  
XX  
XX receptor; mouse; gene therapy; gene; ss.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2763

```

FT      /*tag= a
XX      /product= "Her-2/neu ECD-PD fusion"
XX
XX      WO200212341-A2.
XX
XX      14-FEB-2002.
XX
XX      03-AUG-2001; 2001WO-US24283.
XX
XX      03-AUG-2000; 2000US-0632507.
XX
XX      (CORI-) CORIXA CORP.
XX      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX      Cheever MA, Gheysen D;
XX
XX      WPI; 2002-241743/29.
XX      P-PSDB; AAM51152.
XX
XX      Her-2/neu fusion protein for treating or preventing cancer by eliciting
XX      or enhancing an immune response to the protein, has Her-2/neu
XX      extracellular domain fused to Her-2/neu intracellular or
XX      phosphorylation domain
XX
XX      Disclosure; Fig 23; 141pp; English.
XX
XX      The present sequence is that of cDNA encoding a fusion between
XX      the extracellular domain (ECD) and the phosphorylation domain (PD)
XX      murine Her-2/neu oncoprotein (see AAM51151). The Her-2/neu gene is
XX      amplified and the oncoprotein is overexpressed in a variety of
XX      human cancers, including breast, ovarian, colon, lung and prostate
XX      cancer. Her-2/neu overexpression correlates with a poor prognosis
XX      in breast and ovarian cancers. The invention provides Her-2/neu
XX      fusion proteins, nucleic acids encoding them, viral vectors, and
XX      vaccines comprising the fusion proteins or nucleic acid molecules.
XX      In preferred fusion proteins, the ECD of Her-2/neu is fused to a
XX      Her-2/neu intracellular domain or PD (or its DeltapD fragment). An
XX      immune response to Her-2/neu protein is elicited or enhanced by
XX      administering the fusion protein in the form of a vaccine, or by
XX      transfecting cells of an animal ex vivo with a nucleic acid
XX      encoding the fusion protein, and delivering the transfected cells
XX      to the animal. The fusion proteins, nucleic acids, and isolated
XX      specific T-cells are useful for inhibiting the development of a
XX      cancer, especially breast, ovarian, colon, lung or prostate cancer
XX      in a patient. T cells that specifically react with a Her-2/neu
XX      fusion protein can be used to remove tumour cells from a sample in
XX      order to inhibit the development of cancer in a patient.
XX
XX      Sequence 2763 BP; 571 A; 855 C; 772 G; 565 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.: 1.96e-11 Length: 2763
XX      Score: 147.00 Matches: 27
XX      Percent Similarity: 96.67% Conservative: 2
XX      Best Local Similarity: 90.00% Mismatches: 1
XX      Query Match: 91.88% Indels: 0
XX      DB: 24 Gaps: 0
XX
XX      JUNC_SEQ8_SEQ5 (1-30) x ABA92252 (1-2763)
XX
XX      Qy 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
XX      |||||||.....
XX      Db 1918 GAACGAGGCTGCCAGCAGAGCAGAGAGCCAGCCAGTGCAGGAGGAGGACTTA 1977
XX
XX      Qy 21 GlyProAlaSerProLeuAspSerThrPhe 30
XX      |||||||.....
XX      Db 1978 GGCCCCCAGCCAGCCCATGGAGCAGCACCTTC 2007
XX
XX      RESULT 2
XX      ID ABA92253 standard; cDNA; 2781 BP.
XX
XX      AC ABA92253;

```

```

XX      17-JUN-2002 (first entry)
XX
XX      Her-2/neu extracellular-phosphorylation domain-Tcp0 fusion cDNA.
XX
XX      Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
XX      receptor; Tcp0; mouse; gene therapy; gene; ss.
XX
XX      Chimeric - Mus musculus.
XX      OS Chimeric - Unidentified.
XX
XX      Key Location/Qualifiers
XX      CDS 1..2781
XX      FT /*tag= a
XX      FT /product= "Her-2/neu ECD-PD-Tcp0 fusion"
XX
XX      WO200212341-A2.
XX
XX      14-FEB-2002.
XX
XX      03-AUG-2001; 2001WO-US24283.
XX
XX      03-AUG-2000; 2000US-0632507.
XX
XX      (CORI-) CORIXA CORP.
XX      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX      Cheever MA, Gheysen D;
XX
XX      WPI; 2002-241743/29.
XX      P-PSDB; AAM51153.
XX
XX      Her-2/neu fusion protein for treating or preventing cancer by eliciting
XX      or enhancing an immune response to the protein, has Her-2/neu
XX      extracellular domain fused to Her-2/neu intracellular or
XX      phosphorylation domain
XX
XX      Disclosure; Fig 25; 141pp; English.
XX
XX      The present sequence is that of cDNA encoding a fusion between
XX      the extracellular domain (ECD) and the phosphorylation domain (PD)
XX      murine Her-2/neu oncoprotein (see AAM51151) plus a C-terminal
XX      Tcp0 motif that improves immunogenicity. The Her-2/neu gene is
XX      amplified and the oncoprotein is overexpressed in a variety of
XX      human cancers, including breast, ovarian, colon, lung and prostate
XX      cancer. Her-2/neu overexpression correlates with a poor prognosis
XX      in breast and ovarian cancers. The invention provides Her-2/neu
XX      fusion proteins, nucleic acids encoding them, viral vectors, and
XX      vaccines comprising the fusion proteins or nucleic acid molecules.
XX      In preferred fusion proteins, the ECD of Her-2/neu is fused to a
XX      Her-2/neu intracellular domain or PD (or its DeltapD fragment). An
XX      immune response to Her-2/neu protein is elicited or enhanced by
XX      administering the fusion protein in the form of a vaccine, or by
XX      transfecting cells of an animal ex vivo with a nucleic acid
XX      encoding the fusion protein, and delivering the transfected cells
XX      to the animal. The fusion proteins, nucleic acids, and isolated
XX      specific T-cells are useful for inhibiting the development of a
XX      cancer, especially breast, ovarian, colon, lung or prostate cancer
XX      in a patient. T cells that specifically react with a Her-2/neu
XX      fusion protein can be used to remove tumour cells from a sample in
XX      order to inhibit the development of cancer in a patient.
XX
XX      Sequence 2781 BP; 574 A; 859 C; 779 G; 569 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.: 1.97e-11 Length: 2781
XX      Score: 147.00 Matches: 27
XX      Percent Similarity: 96.67% Conservative: 2
XX      Best Local Similarity: 90.00% Mismatches: 1
XX      Query Match: 91.88% Indels: 0
XX      DB: 24 Gaps: 0
XX
XX      JUNC_SEQ8_SEQ5 (1-30) x ABA92253 (1-2781)

```



QY 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20  
 Db 1918 GAACGAGGCTGCCAGCAGACAGACAGAGCCAGCGTCTCTCAGACGAGGACTTA 1977

QY 21 GlyProAlaSerProLeuAspSerThrPhe 30  
 Db 1978 GGCCCTCCAGCCCATGAGCAGCACCTTC 2007

RESULT 3  
 AAF21778  
 ID AAF21778 standard; DNA; 1115 BP.  
 XX  
 AC AAF21778;  
 XX  
 DT 27-MAR-2001 (first entry)  
 XX  
 DE Human breast and ovarian cancer associated antigen gene SEQ ID 165.  
 XX  
 KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;  
 KW nontropic; neuroprotective; antiviral; anti-allergic; hepatotropic;  
 KW anti-diabetic; anti-inflammatory; anti-ulcer; vulnery; anticonvulsant;  
 KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;  
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;  
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;  
 KW cardiovascular disorder; wound healing; neurological disease; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200055173-A1.  
 PN  
 XX  
 PD 21-SEP-2000.  
 XX  
 XX 08-MAR-2000; 2000WO-US05881.  
 XX  
 PR 12-MAR-1999; 99US-0124270.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Rosen CA, Ruben SM;  
 PI  
 XX WPI; 2000-611515/58.  
 DR P-PSDB; AAB58875.  
 XX  
 XX New human breast and ovarian cancer associated gene sequences and the  
 PT polypeptides encoded by these genes, useful in the prevention,  
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular  
 PT disorders and neurological diseases -  
 XX  
 PS Claim 1; Page 604; 1299pp; English.  
 XX  
 CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are  
 CC associated with breast and ovarian cancer. Included in the invention are  
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
 CC isolation and characterisation of the DNA and protein sequences of the  
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist  
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;  
 CC nontropic; neuroprotective; antiviral; anti-allergic; hepatotropic;  
 CC anti-diabetic; anti-inflammatory; anti-ulcer; vulnery; anticonvulsant;  
 CC antibacterial; antifungal; antiparasitic and cardiant activity. The  
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,  
 CC particularly breast and ovarian cancer. The nucleic acid sequences,  
 CC proteins, agonists and antagonists may also be used in the diagnosis,  
 CC prevention and treatment of immune disorders e.g. Addison's disease,  
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
 CC arthritis and ulcerative colitis; cardiovascular disorders such as  
 CC myocardial ischaemias; wound healing; neurological diseases such as  
 CC cerebral anoxia and epilepsy; and infectious diseases.  
 XX  
 SQ Sequence 1115 BP; 210 A; 336 C; 338 G; 222 T; 9 other;

Alignment Scores:  
 Pred. No.: 0.0139 Length: 1115  
 Score: 83.00 Matches: 16  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 51.88% Indels: 0  
 DB: 21 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x AAF21778 (1-1115)  
 QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 Db 446 GTGGTCATCCAGAAATGAGGACTTGGCCAGCAGTCCCTTGGACAGCACCTTC 499

RESULT 4  
 AAD32746  
 ID AAD32746 standard; cDNA; 1755 BP.  
 XX  
 AC AAD32746;  
 XX  
 DT 01-JUL-2002 (first entry)  
 XX  
 DE Human cDNA for the clone HICD\_native\_coding\_region.  
 XX  
 KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;  
 KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1755  
 FT /\*tag= a  
 FT /product= "Human protein encoded by cDNA for the clone  
 FT HICD\_native\_coding\_region"  
 FT /transl\_except= (pos:1741..1752, aa:Leu-Glu)  
 FT /note= "CDS does not include stop codon"  
 FT /partial  
 XX  
 XX WO200214503-A2.  
 PN  
 XX  
 PD 21-FEB-2002.  
 XX  
 XX 14-AUG-2001; 2001WO-US41733.  
 XX  
 PR 14-AUG-2000; 2000US-225152P.  
 PR 28-SEP-2000; 2000US-236428P.  
 PR 21-FEB-2001; 2001US-270520P.  
 XX  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;  
 PI McNeill PD, Vedvick TS;  
 XX  
 DR WPI; 2002-280758/32.  
 DR P-PSDB; AAE20483.  
 XX  
 XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,  
 PT prevention and diagnosis of cancer, preferably breast cancer -  
 XX  
 XX Example 5; Page 118-119; 129pp; English.  
 XX  
 CC The invention relates to an isolated Her-2/Neu polypeptide composition  
 CC effective for eliciting an immune response. The invention is useful for  
 CC eliciting an immune response in a patient, where the patient is human  
 CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.  
 CC The composition is useful for the therapy and diagnosis of cancer,  
 CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine  
 CC and other compositions for the diagnosis, prevention and treatment of  
 CC human malignancies, for stimulating and/or expanding T cells specific for  
 CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a  
 CC patient. The invention is useful for stimulating a T cell response in a  
 CC human patient, as probe or primer for nucleic acid hybridisation, to



XX Example 5; Page 119; 129pp; English.

XX The invention relates to an isolated Her-2/Neu polypeptide composition

CC effective for eliciting an immune response. The invention is useful for

CC eliciting an immune response in a patient, where the patient is human

CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.

CC The composition is useful for the therapy and diagnosis of cancer,

CC and other compositions, in pharmaceutical compositions, e.g., vaccine

CC human malignancies, for stimulating and/or expanding T cells specific for

CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a

CC patient. The invention is useful for stimulating a T cell response in a

CC human patient, as probe or primer for nucleic acid hybridisation, to

CC selectively form duplex molecules with complementary stretches of the

CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full

CC length gene from a suitable library, and to direct expression of a

CC polypeptide in appropriate host cells. The composition is useful in

CC prophylactic or therapeutic applications and for the treatment of cancer,

CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-

CC associated malignancies. The invention is useful in gene therapy. The

CC present sequence is human cDNA for the clone HICD\_in\_pPDM\_coding\_

XX sequence.

XX Sequence 1773 BP; 383 A; 528 C; 530 G; 332 T; 0 other;

Alignment Scores:

Pred. No.: 0.0242 Length: 1773

Score: 83.00 Matches: 16

Percent Similarity: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 2

Query Match: 51.88% Indels: 0

DB: 24 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x AAD32747 (1-1773)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 961 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 1014

RESULT 7

AAD32745

ID AAD32745 standard; cDNA; 1806 BP.

XX AAD32745;

01-JUL-2002 (first entry)

XX Human cDNA for the clone HICD\_plus\_8\_HIS.

XX Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;

KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1803

FT /\*tag= a

FT /product= "Human protein encoded by cDNA for the clone

FT HICD\_plus\_8\_HIS"

FT /transl\_except= (pos:1543..1545, aa:Pro)

XX WO200214503-A2.

XX 21-FEB-2002.

XX 14-AUG-2001; 2001WO-US41733.

XX 14-AUG-2000; 2000US-225152P.

PR 28-SEP-2000; 2000US-236428P.

PR 21-FEB-2001; 2001US-270520P.

XX (CORI-) CORIXA CORP.

XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;

PI McNeill PD, Vedvick TS;

XX WPI: 2002-280758/32.

DR P-PSDB; AAE20482.

XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,

PT prevention and diagnosis of cancer, preferably breast cancer

XX Example 5; Page 118; 129pp; English.

XX The invention relates to an isolated Her-2/Neu polypeptide composition

CC effective for eliciting an immune response. The invention is useful for

CC eliciting an immune response in a patient, where the patient is human

CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.

CC The composition is useful for the therapy and diagnosis of cancer,

CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine

CC and other compositions, for stimulating and/or expanding T cells specific for

CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a

CC patient. The invention is useful for stimulating a T cell response in a

CC human patient, as probe or primer for nucleic acid hybridisation, to

CC selectively form duplex molecules with complementary stretches of the

CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full

CC length gene from a suitable library, and to direct expression of a

CC polypeptide in appropriate host cells. The composition is useful in

CC prophylactic or therapeutic applications and for the treatment of cancer,

CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-

CC associated malignancies. The invention is useful in gene therapy. The

CC present sequence is human cDNA for the clone HICD\_plus\_8\_HIS.

XX Sequence 1806 BP; 391 A; 530 C; 544 G; 341 T; 0 other;

Alignment Scores:

Pred. No.: 0.0248 Length: 1806

Score: 83.00 Matches: 16

Percent Similarity: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 2

Query Match: 51.88% Indels: 0

DB: 24 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x AAD32745 (1-1806)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 994 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 1047

RESULT 8

AAA89736

ID AAA89736 standard; DNA; 3600 BP.

XX AAA89736;

12-JAN-2001 (first entry)

XX Human HER-2/neu coding sequence.

XX Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;

KW breast cancer; prostate cancer; ovarian cancer; lung cancer;

KW colon cancer; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..3600

FT /\*tag= a

FT /product= "HER-2/neu protein"

XX WO200044899-A1.

XX 03-AUG-2000.

PF 28-JAN-2000; 2000WO-US02164.  
XX  
PR 29-JAN-1999; 99US-0117976.  
XX  
PA (CORI-) CORIXA CORP.  
PA (SMIK ) SMITHKLINE BEECHAM.  
XX  
PI Cheever MA, Gheysen D;  
XX  
XX WPI; 2000-505976/45.  
DR P-PSDB; AAB21198, AAB21208.  
DR  
XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins  
PT useful for vaccinating against breast, ovarian, colon, lung and  
PT prostate cancers -  
XX  
PS Disclosure: Fig 15; 128pp; English.  
XX  
CC The present sequence encodes the human HER-2/neu protein. HER-2/neu is  
CC a member of the tyrosine kinase family of receptor-like glycoproteins  
CC and shows homology to the epidermal growth factor receptor (EGFR). It  
CC probably plays a part in cell growth and/or differentiation. The  
CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising  
CC a HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation  
CC domain may be used to treat or prevent cancer by eliciting or  
CC enhancing an immune response to the HER-2/neu protein. It may be used  
CC to treat malignancies such as breast, ovarian, colon, lung and  
CC prostate cancers, and may be used as an antigen to vaccinate against  
CC these neoplasias.  
XX  
SQ Sequence 3600 BP; 723 A; 1108 C; 1075 G; 694 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 0.0566 Length: 3600  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 21 Gaps: 0  
  
JUNC\_SEQ8\_SEQ5 (1-30) x AAB89736 (1-3600)  
  
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACACACCTTC 3012  
  
RESULT 9  
ABK86207  
ID ABR86207 standard; cDNA; 3678 BP.  
XX  
AC ABR86207;  
XX  
XX 24-SEP-2002 (first entry)  
XX  
XX cDNA encoding human breast cancer antigen, Her2 variant.  
DE  
DE Human; Her2; cytostatic; antiviral; immunostimulant;  
KW cell-mediated immune response; tumour; breast cancer;  
KW virus infection; prostate cancer; colorectal cancer; pancreatic cancer;  
KW lymphoma; leukaemia; hepatitisvirus; lentivirus; herpesvirus;  
KW human immunodeficiency virus; HIV; flavivirus; pestivirus; gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
FT CDS 7..3678  
FT /tag= a  
FT /product= "Breast cancer antigen Her2 variant"  
XX  
XX WO2002040059-A2.  
XX  
XX 23-MAY-2002.  
PD  
XX

PF 01-NOV-2001; 2001WO-US45636.  
XX  
PR 01-NOV-2000; 2000US-0704232.  
XX  
PA (AMBI-) AMERICAN FOUND BIOLOGICAL RES INC.  
PA (MINC/) MINCHEFF M S.  
PA (LOUK/) LOUKINOV D I.  
PA (ZOUB/) ZOUBAK S.  
XX  
XX Mincheff MS, Loukinov DI, Zoubak S;  
PI WPI; 2002-527524/56.  
XX P-PSDB; AAU98923.  
DR  
DR Inducing a cell-mediated immune response against a target antigen,  
PT reducing undesired cells and stimulating presentation of an antigen by  
PT a cell, comprises administering a polynucleotide encoding a variant of  
PT an antigen -  
XX  
XX Disclosure: Page 128-134; 146pp; English.  
XX  
CC The invention relates to a method of inducing a cell-mediated immune  
CC response against a cell comprising a target antigen (I) in a subject,  
CC treating a subject having undesired cells, for example tumour cells  
CC or virally infected cells (C), reducing the number of (C) in a subject,  
CC and stimulating presentation of (I) by a cell. This is done by  
CC administering a polynucleotide (II) encoding a variant of (I), so that  
CC (II) expressed in a cell and cell-mediated immune response is induced.  
CC The method can be used to treat prostate cancer, breast cancer,  
CC colorectal cancer and pancreatic cancer, as well as lymphomas and  
CC leukemias. The method is also useful in treating chronic viral  
CC infections such as those caused by hepatitisviruses, lentiviruses  
CC (including human immunodeficiency virus (HIV)), herpesviruses and the  
CC flaviviruses and pestiviruses. The present sequence represents the coding  
CC sequence of human breast cancer antigen, Her2 variant, used as a target  
XX antigen in the method of the invention.  
XX  
SQ Sequence 3678 BP; 746 A; 1137 C; 1091 G; 704 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 0.0581 Length: 3678  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 24 Gaps: 0  
  
JUNC\_SEQ8\_SEQ5 (1-30) x ABR86207 (1-3678)  
  
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 2869 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACACACCTTC 2922  
  
RESULT 10  
AAT40739  
ID AAT40739 standard; cDNA; 3768 BP.  
XX  
XX AAT40739;  
AC AAT40739;  
XX  
XX 01-JAN-1997 (first entry)  
DT  
XX HER-2/neu oncogene.  
DE  
DE  
XX  
XX HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;  
KW breast cancer; ovary cancer; colon cancer; lung cancer;  
KW prostate cancer; genetic immunisation; tumour; vaccine; vector;  
KW ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FT CDS 1..3765  
FT /\*tag= b  
FT

```

FT      /note= "nucleotides 2026-3765 (claim 1) code for
FT      HER-2/neu intracellular domain"
PN      WO9630514-A1.
XX
XX      03-OCT-1996.
XX
XX      28-MAR-1996; 96WO-US01689.
XX
XX      31-MAR-1995; 95US-0414417.
XX      (UNIW ) UNIV WASHINGTON.
PA      Cheever MA, Disis ML;
XX      WPI; 1996-455361/45.
XX      P-PSDB; AAW01111.
DR
XX      DNA encoding HER-2-neu poly:peptide(s) - used for prevention or
PT      treatment of malignancies with which the HER-2/neu oncogene is
PT      associated
XX
XX      Claim 1; Page 49-56; 71pp; English.
XX
XX      Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or
CC      c-erbB2) protein (AAW01111). The oncogene is overexpressed in various
CC      cancers, including breast, ovarian, colon, lung and prostate, and
CC      appears to induce malignancies through quantitative mechanisms that
CC      result from increased or deregulated expression of an essentially
CC      normal gene product. Nucleotides 2026-3765 of the cDNA sequence
CC      code for the intracellular domain (lys676-Val1255) of the HER-2/neu
CC      protein, which is useful for immunisation against malignancy.
CC      Nucleic acids can be used to direct expression of the intracellular
CC      domain in transformed host cells, or are used, alone or in a viral
CC      vector, for genetic immunisation of an animal.
XX
SQ      Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 17 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AAT40739 (1-3768)
QY      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
      ||| |||||||||||||||||||||||||||||||||||||||||||||||
Db      2959 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 3012

RESULT 11
AAAX01912
ID      AAX01912 standard; DNA; 3768 BP.
XX
XX      AAX01912;
AC
XX      21-APR-1999 (first entry)
DT
XX      Human HER-2/neu oncogene DNA.
DE
XX      HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
KW      malignancy; treatment; tumour; ss.
KW
XX      Homo sapiens.
OS
XX      Key Location/Qualifiers
FH      1..3768
FT      CDS
FT      /tag= a
FT      /product= "HER-2/neu"
FT      /note= "oncogene"
FT      2026..3765
FT      misc_feature

```

```

FT      /*tag= b
FT      /note= "region which elicits immune response"
PN      US5869445-A.
XX
XX      09-FEB-1999.
XX
XX      01-APR-1996; 96US-0625101.
XX
XX      01-APR-1996; 96US-0625101.
PR      17-MAR-1993; 93US-0033644.
PR      12-AUG-1993; 93US-0106112.
PR      31-MAR-1995; 95US-0414417.
XX
XX      (UNIW ) UNIV WASHINGTON.
PA
XX      Cheever MA, Disis ML;
PI
XX      WPI; 1999-152835/13.
DR      P-PSDB; AAW92406.
XX
XX      Use of HER-2/neu polypeptides - for eliciting an immune response to
PT      an HER-2/neu associated malignancy, particularly for treating or
PT      preventing tumours
XX
XX      Claim 1a; Column 23-32; 26pp; English.
XX
XX      This sequence encodes the human HER-2/neu oncogene protein. A fragment
CC      of this protein is used in a method for eliciting or enhancing an immune
CC      response to HER-2/neu protein. The polypeptide can stimulate T cells and
CC      B cells to produce an immune response to the HER-2/neu protein. The
CC      method can be used for immunisation against a malignancy in which the
CC      HER-2/neu oncogene is associated and in the treatment of an existing
CC      tumour, or to prevent tumour occurrence or reoccurrence.
XX
SQ      Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 20 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AAX01912 (1-3768)
QY      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
      ||| |||||||||||||||||||||||||||||||||||||||||||||||
Db      2959 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 3012

RESULT 12
AAAX09455
ID      AAX09455 standard; DNA; 3768 BP.
XX
XX      AAX09455;
AC
XX      10-AUG-2000 (first entry)
DT
XX      Human heregulin 2 (Her2) coding sequence.
DE
XX      Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
KW      self-protein; cell-associated peptide antigen; foreign epitope;
KW      cancer; breast cancer; prostate cancer; ss.
KW
XX      Homo sapiens.
OS
XX      WO200020027-A2.
PN
XX      13-APR-2000.
PD
XX      05-OCT-1999; 99WO-DK00525.
PF
XX

```

```
PR 05-OCT-1998; 98DK-0001261.
XX 20-OCT-1998; 98US-0105011.
XX (MEBI-) M & E BIOTECH AS.
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX WPI; 2000-349917/30.
DR P-PSDB; AAY92620.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
XX Claim 62; Page 187-193; 220pp; English.
XX
XX The claims detail a method for inducing immune responses against weakly
CC immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (i.e. self-proteins), for example, human
CC prostate specific membrane antigen (PSM), heregulin 2 (Her2), and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively.
XX
XX Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 21 Gaps: 0
JUNC_SEQ8_SEQ5 (1-30) x AAA09455 (1-3768)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 3012
RESULT 13
AAH23392
ID AAH23392 standard; DNA; 3768 BP.
XX
XX AAH23392;
XX
XX 25-SEP-2001 (first entry)
XX Human HER-2/neu protein encoding DNA.
XX
XX Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
KW oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 1..3768
FT /*tag= a
FT /product= "HER-2/neu protein"
XX
XX WO200153463-A2.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US01850.
```

```
XX 21-JAN-2000; 2000US-0177545.
XX (CORI-) CORIXA CORP.
XX Cheever MA, Hand-Zimmermann S;
PI WPI; 2001-476112/51.
XX P-PSDB; AAB85458.
XX
XX New antigen-presenting cells, useful as vaccines for eliciting or
PT enhancing an immune response to HER-2/neu protein, particularly useful
PT for treating or preventing cancer, e.g. breast cancer
XX
XX Claim 1; Page 41-46; 49pp; English.
XX
XX The invention provides an isolated antigen-presenting cell, which
CC expresses at least an immunogenic portion of a polypeptide that produces
CC an immune response to HER-2/neu protein. The antigen-presenting cells are
CC useful as vaccines for eliciting or enhancing an immune response to
CC HER-2/neu protein, particularly in treating or preventing malignancies in
CC which the HER-2/neu oncogene is associated. Specifically, these are
CC useful for treating or preventing cancer, e.g. breast cancer, ovarian,
CC colon, lung or prostate cancers. The present sequence represents a DNA
CC encoding the human HER-2/neu protein (also known as p185 or c-erbB2).
XX
XX Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 22 Gaps: 0
JUNC_SEQ8_SEQ5 (1-30) x AAH23392 (1-3768)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 3012
RESULT 14
ABZ35744
ID ABZ35744 standard; DNA; 3768 BP.
XX
XX ABZ35744;
XX
XX 07-FEB-2003 (first entry)
XX Human ERBB2 polynucleotide SEQ ID NO 52.
XX
XX Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
KW protozoacide; gene expression; antisense; tumour; infection; plasmodium;
KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
KW Hepatitis C virus; human papilloma virus; gene; ds.
XX
XX Homo sapiens.
XX
XX DE10100588-A1.
XX
XX 18-JUL-2002.
XX
XX 09-JAN-2001; 2001DE-1000588.
XX
XX 09-JAN-2001; 2001DE-1000588.
XX (RIBO-) RIBOPHARMA AG.
XX
XX Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
XX WPI; 2002-683450/74.
XX
```

PT Inhibiting expression of target genes, useful e.g. for treating tumors,  
PT by introducing into cells two double-stranded RNAs that are  
PT complementary to the target  
XX  
XX Claim 13; Page 38-39; 100pp; German.  
XX  
CC The invention relates to inhibiting expression of a target gene in a cell  
CC by introducing at least two oligoribonucleotides (dsRNAI and II), both  
CC with a double-stranded (ds) structure of at most 49 sequential nucleotide  
CC pairs. At least part of one strand (S1, S2) of the ds structures in each  
CC of dsRNAI and II are complementary to regions in the target gene. The  
CC method uses antisense inhibition of gene expression using double stranded  
CC RNA inhibition (RNAi). The method is particularly used to treat tumours  
CC or infections, especially by Plasmodium or viruses/viroids (pathogenic on  
CC humans, animals or plants). The method provides more effective inhibition  
CC of expression than known methods using a single dsRNA, even at very low  
CC concentrations. When dsRNA has at least one unpaired nucleotide at the  
CC end, stability (and thus effective concentration in the cell) is  
CC improved and efficiency can be increased further by pretreating the cells  
CC with interferon. The present sequence is that of a target DNA of the  
CC invention.  
XX  
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:  
Pred. No.: 0.0598 Length: 3768  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 24 Gaps: 0  
JUNC\_SEQ8\_SEQ5 (1-30) x ABZ35744 (1-3768)  
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 3012

RESULT 15  
ABX09987  
ID ABX09987 standard; DNA; 3768 BP.  
XX  
AC ABX09987;  
XX  
DT 23-JAN-2003 (first entry)  
XX  
DE Human ERBB2 DNA fragment SEQ ID 52.  
XX  
KW Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;  
KW prion; inhibition; human; ds.  
XX  
OS Homo sapiens.  
XX  
PN DE10100587-C1.  
XX  
PD 21-NOV-2002.  
XX  
PF 09-JAN-2001; 2001DE-1000587.  
XX  
PR 09-JAN-2001; 2001DE-1000587.  
XX  
PA (RIBO-) RIBOPHARMA AG.  
XX  
PI Kreutzer R, Limmer S, Rost S, Hadwiger P;  
XX  
XX WPI; 2002-742209/81.  
DR  
XX  
XX Inhibiting expression of target genes, e.g. oncogenes, in cells, by  
PT introduction of complementary double-stranded oligoribonucleotide,  
PT after treating the cell with interferon  
XX  
XX Disclosure; Page 43-44; 98pp; German.  
XX

CC This invention describes a novel method for inhibiting expression of a  
CC target gene by introducing into the cell that contains the target gene  
CC at least one oligoribonucleotide (dsRNAI) that has a double-stranded  
CC (ds) structure of not more than 49 consecutive nucleotides (nt), where  
CC at least a segment of one strand of the ds structure is complementary  
CC with the target gene and the cells are treated with interferon before  
CC introduction of dsRNAI. The method is used to inhibit expression of  
CC target genes, particularly oncogenes, cytokine genes, Id (not defined)  
CC protein genes; developmental or prion genes, or genes expressed in  
CC pathogenic organisms (particularly plasmodia) or in viruses or viroids  
CC (pathogenic in humans, animals or plants). Treating the cells with  
CC interferon greatly increases the extent to which dsRNA can inhibit  
CC expression of the target genes, and the effect is even greater when dsRNA  
CC are modified to increase their stability. ABX09936-ABX10075 represent  
XX gene fragments used to illustrate the method of the invention.

SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;  
Alignment Scores:  
Pred. No.: 0.0598 Length: 3768  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 24 Gaps: 0  
JUNC\_SEQ8\_SEQ5 (1-30) x ABX09987 (1-3768)  
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 3012

Search completed: October 15, 2003, 23:48:52  
Job time : 32.4941 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:09:13 ; Search time 432.026 Seconds  
(without alignments)  
2840.777 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ5  
Perfect score: 160  
Sequence: 1 ERGCPAEQRASPVTFQNEGLGPASPLDSTF 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV-xlp  
-MODEL-frame+p2n.model  
-Q/cgn2\_1/USPTO\_spool\_p/HOLLERAN480/runat\_15102003\_131912\_20336/app\_query.fasta\_1.4685  
-DB-GenEmbl -FAST=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS-human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480.GCGN\_1\_1\_2724@runat\_15102003\_131912\_20336 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.or.\*
- 21: em.ov.\*
- 22: em.pat.\*
- 23: em.ph.\*
- 24: em.pl.\*
- 25: em.ro.\*
- 26: em.sts.\*
- 27: em.un.\*
- 28: em.un.\*

- 29: em.vi.\*
- 30: em.htg\_hum.\*
- 31: em.htg\_inv.\*
- 32: em.htg\_other.\*
- 33: em.htg\_mus.\*
- 34: em.htg\_pln.\*
- 35: em.htg\_rod.\*
- 36: em.htg\_mam.\*
- 37: em.htg\_vrt.\*
- 38: em.sy.\*
- 39: em.htgo\_hum.\*
- 40: em.htgo\_mus.\*
- 41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	147	91.9	2763	6	AX380942	AX380942 Sequence
2	147	91.9	2781	6	AX380944	AX380944 Sequence
3	83	51.9	1755	6	AX384609	AX384609 Sequence
4	83	51.9	1767	6	AX384607	AX384607 Sequence
5	83	51.9	1773	6	AX384610	AX384610 Sequence
6	83	51.9	1806	6	AX384608	AX384608 Sequence
7	83	51.9	3678	6	AX505114	AX505114 Sequence
8	83	51.9	3768	6	AR034479	AR034479 Sequence
9	83	51.9	3768	6	AX060704	AX060704 Sequence
10	83	51.9	3768	6	AX201817	AX201817 Sequence
11	83	51.9	3768	6	AX380923	AX380923 Sequence
12	83	51.9	3768	6	AX384604	AX384604 Sequence
13	83	51.9	3768	6	AX465456	AX465456 Sequence
14	83	51.9	3768	6	AX467229	AX467229 Sequence
15	83	51.9	3768	6	AX481438	AX481438 Sequence
16	83	51.9	3780	4	AB008451	AB008451 Canis fam
17	83	51.9	4473	6	AR080259	AR080259 Sequence
18	83	51.9	4473	6	AR167390	AR167390 Sequence
19	83	51.9	4473	9	HSEB2R	X03363 Human c-erb
20	83	51.9	4530	6	AR202597	AR202597 Sequence
21	83	51.9	4530	6	AR283481	AR283481 Sequence
22	83	51.9	4530	6	AX282577	AX282577 Sequence
23	83	51.9	4530	6	AX587649	AX587649 Sequence
24	83	51.9	4530	6	AX644071	AX644071 Sequence
25	83	51.9	4530	6	BD005474	BD005474 Cellular
26	83	51.9	4530	6	I21124	I21124 Sequence 9
27	83	51.9	4530	6	I59745	I59745 Sequence 9
28	83	51.9	4530	9	HUMHER2A	M11730 Human tyros
29	83	51.9	9274	6	AX060703	AX060703 Sequence
30	81	50.6	3771	6	AX189662	AX189662 Sequence
31	81	50.6	3771	6	AX380925	AX380925 Sequence
32	81	50.6	3955	6	AX380924	AX380924 Sequence
33	81	50.6	3955	6	I21129	I21129 Sequence 14
34	81	50.6	3955	6	I59750	I59750 Sequence 14
35	81	50.6	3955	10	RNNEUR	X03362 Rat mRNA fo
36	81	50.6	4694	10	BC046811	BC046811 Mus muscu
37	81	50.6	4695	10	BC053078	BC053078 Mus muscu
38	81	50.6	4727	10	AY116182	AY116182 Rattus no
39	80	50.0	4062	10	HAMNEU	D16295 Mesocricetu
40	79	49.4	13450	9	AB096612	AB096612 Homo sapi
41	79	49.4	16572	9	AB096614	AB096614 Homo sapi
42	79	49.4	20271	9	AB096613	AB096613 Homo sapi
43	79	49.4	30837	9	AY208911	AY208911 Homo sapi
44	79	49.4	161815	9	AC079199	AC079199 Homo sapi
45	79	49.4	168585	9	AC040933	AC040933 Homo sapi

ALIGNMENTS

RESULT 1

```

AX380942
LOCUS AX380942 2763 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 28 from Patent WO0212341.
ACCESSION AX380942
VERSION AX380942.1 GI:19575786
KEYWORDS
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
AUTHORS Cheever, M.A. and Gheysen, D.
TITLE Her-2/neu fusion proteins
JOURNAL Patent: WO 0212341-A 28 14-FEB-2002;
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
source
Location/Qualifiers
1..2763
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mouse ECD-PD fusion protein cDNA"
BASE COUNT 571 a 855 c 772 g 565 t
ORIGIN
Alignment Scores:
Pred. No.: 3.46e-11 Length: 2763
Score: 147.00 Matches: 27
Percent Similarity: 96.67% Conservative: 2
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 91.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AX380942 (1-2763)
Qy 1 GluArgGlyCysProAlaGluInArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCAGCAGACGAGAGCCAGCCAGTGTCTCAGAACGAGGACTTA 1977
Qy 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 1978 GGCCCCCTCCAGCCCATGGACAGCACCTTC 2007
RESULT 2
AX380944
LOCUS AX380944 2781 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 30 from Patent WO0212341.
ACCESSION AX380944
VERSION AX380944.1 GI:19575787
KEYWORDS
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
AUTHORS Cheever, M.A. and Gheysen, D.
TITLE Her-2/neu fusion proteins
JOURNAL Patent: WO 0212341-A 30 14-FEB-2002;
CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
source
Location/Qualifiers
1..2781
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mouse ECD-PD-Tcp0 fusion protein cDNA"
BASE COUNT 574 a 859 c 779 g 569 t
ORIGIN
Alignment Scores:
Pred. No.: 3.49e-11 Length: 2781
Score: 147.00 Matches: 27
Percent Similarity: 96.67% Conservative: 2
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 91.88% Indels: 0
DB: 6 Gaps: 0

```

```

JUNC_SEQ8_SEQ5 (1-30) x AX380944 (1-2781)
Qy 1 GluArgGlyCysProAlaGluInArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCAGCAGACGAGAGCCAGCCAGTGTCTCAGAACGAGGACTTA 1977
Qy 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 1978 GGCCCCCTCCAGCCCATGGACAGCACCTTC 2007
RESULT 3
AX384609
LOCUS AX384609 1755 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 6 from Patent WO0214503.
ACCESSION AX384609
VERSION AX384609.1 GI:19577810
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
JOURNAL McNeill, P.D. and Vedvick, T.S.
Patent: WO 0214503-A 6 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..1755
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 376 a 517 c 531 g 331 t
ORIGIN
Alignment Scores:
Pred. No.: 0.0303 Length: 1755
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ5 (1-30) x AX384609 (1-1755)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 937 GTGGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 990
RESULT 4
AX384607
LOCUS AX384607 1767 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 4 from Patent WO0214503.
ACCESSION AX384607
VERSION AX384607.1 GI:19577808
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
JOURNAL McNeill, P.D. and Vedvick, T.S.
Patent: WO 0214503-A 4 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..1767
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

BASE COUNT 381 a 521 c 529 g 336 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.0305 Length: 1767  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 6 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x AX384607 (1-1767)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 937 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGCTCCCTTGGACAGCACCTTC 990

RESULT 5  
AX384610  
LOCUS AX384610 1773 bp DNA linear PAT 19-MAR-2002  
DEFINITION Sequence 7 from Patent WO0214503.  
ACCESSION AX384610  
VERSION AX384610.1 GI:19577811  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 McNeill, P.D. and Vedvick, T.S.

TITLE Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,  
Compositions and methods for the therapy and diagnosis of  
her-2/neu-associated malignancies  
JOURNAL Patent: WO 0214503-A 7 21-FEB-2002;  
CORIXA CORPORATION (US)

FEATURES  
source  
1. 1773  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 383 a 528 c 530 g 332 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.0306 Length: 1773  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 6 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x AX384610 (1-1773)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 961 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGCTCCCTTGGACAGCACCTTC 1014

RESULT 6  
AX384608  
LOCUS AX384608 1806 bp DNA linear PAT 19-MAR-2002  
DEFINITION Sequence 5 from Patent WO0214503.  
ACCESSION AX384608  
VERSION AX384608.1 GI:19577809  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,  
McNeill, P.D. and Vedvick, T.S.

TITLE Compositions and methods for the therapy and diagnosis of  
her-2/neu-associated malignancies

JOURNAL Patent: WO 0214503-A 5 21-FEB-2002;  
CORIXA CORPORATION (US)

FEATURES  
source  
1. 1806  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 391 a 530 c 544 g 341 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.0312 Length: 1806  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 6 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x AX384608 (1-1806)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 994 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGCTCCCTTGGACAGCACCTTC 1047

RESULT 7  
AX505114  
LOCUS AX505114 3678 bp DNA linear PAT 27-SEP-2002  
DEFINITION Sequence 7 from Patent WO0240059.  
ACCESSION AX505114  
VERSION AX505114.1 GI:23386421  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 Mincheff, M.S., Loukinov, D.I. and Zoubak, S.  
Methods and compositions for inducing cell-mediated immune  
responses  
JOURNAL Patent: WO 0240059-A 7 23-MAY-2002;  
American Foundation for Biological Research Inc. (US); Mincheff,  
Milcho S. (US); Loukinov, Dmitri I. (US); Zoubak, Serguei (US)

FEATURES  
source  
1. 3678  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
7. 3678  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD48804.1"  
/db\_xref="GI:23386422"

/translation="MRLPASPEHLDMRLHLYQGVVQGNLELYLPTNASLSFLQD  
IOEVGYVLIAHNOVROVPLQRLIVRGTOLEFDNYALAVLDNGDLNNTTPVTGASP  
GGLRELQRLSLEILKGVLIORNPOLCYODTILWKDIFHKNNOLATLIDTNSRAC  
GGSPMCKGSRCEGSESSDCQSLRTVPCVGCARCKGPLEPTDCCHCEQCAAGCTPKHS  
DCLACLFHNSGICELHCPALVTDTFESMPAPEPTFGGACCTACCTNYLSTDV  
GCTFLVCLPHNQEVTAEDGTQCEKSKPCARVCYGLGMEHLRVRVAVTSANTIOEFAG  
CKTFGLAFLEPESDGPASNTAPLOEQVFEITLLEETIGYLIISAWPDSLPLSV  
FQNLVIRGRILHNGAYSLTQGLIGISWLGRLSRLGSLGIALIHHNTHLCFVHTVPM  
DOLFRNPHOALLHNGARPEDECGEGLACHOLCARHCWGPGTQCVCNQSLRGQEC  
VEECRVQLGIPREYVNAHCLCPHCEQCPQNGSVTCFGEADOCVACAHYKDPFCVA  
RCPGVRPDLISYMPFIWKFPEEGACQPCPINCTHSCVDLDKGCAPQASPLTSIVS  
AGVGILLVVLGVVFGILIKRRQOKIRKTYMRRLLQETLELVEPLTPSGAMPNQHMRI  
LKETELRKVKVLSGAFCTYGIWIPDGENVKIPVAKVLRNTSPKANKEILDEAY  
VMAGVSPYSVRLGICLTSTVOLTPOLMPYGLDHLVHNRNRLGSDQLLWCMQIA  
KMSYLEEDVRLVHDLAARNVVKSPNHVKITDFGLARLLDIDTEVHAGGVKPIKW  
MALESILRRRTHQSDVNSYGVTWELMTFGAKPYDGIIPAREIPDLLEKGERLPPPI  
CTIDVIMLVKCMWIDSECRFRELVESEFSEMRARDPQRFVITONEDLGPASPLSTFF  
YRSLEDDDDGDLVDABEYLVPOQGFPCPDAPGAGGMVHHRSSSTRSGGDLTLG  
LEFSEEPASPLAPSEAGSDVFDGLGMAAKGLQSLTHDPSPILQRYSEDTPL  
PSETDGYVAPLTCSPQPEYVNPQDVRPQPPSPREGPLPAARPAAGATLERAKTILSPGN  
GVVKDYFAFGGAVENPEYLTPOGGCAAPQPPPPAFSPAFDNLVYWDQDPPERGAPST

BASE COUNT 746 a 1137 c 1091 g 704 t  
ORIGIN

FKGTPTAENPEYLGLDVPV"

Alignment Scores: 0.0666 Length: 3678  
Pred. No.: 83.00 Matches: 16  
Score: 83.00  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 6 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x AX505114 (1-3678)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 2869 GTGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 2922

RESULT 8  
AR034479  
LOCUS AR034479 3768 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 1 from patent US 5869445.  
ACCESSION AR034479  
VERSION AR034479.1 GI:5950084

KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 3768)  
AUTHORS Cheever, M.A. and Disis, M.L.  
TITLE Methods for eliciting or enhancing reactivity to HER-2/neu protein  
JOURNAL Patent: US 5869445-A 1 09-FEB-1999;  
FEATURES Location/Qualifiers  
1..3768

BASE COUNT 759 a 1171 c 1119 g 719 t  
ORIGIN

/organism="unknown"

Alignment Scores: 0.0683 Length: 3768  
Pred. No.: 83.00 Matches: 16  
Score: 83.00  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 6 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x AR034479 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 2959 GTGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 3012

RESULT 9  
AX060704  
LOCUS AX060704 3768 bp DNA linear PAT 22-JAN-2001  
DEFINITION Sequence 2 from Patent WO0100244.  
ACCESSION AX060704  
VERSION AX060704.1 GI:12406101

KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Erickson, S. and Schwall, R.  
TITLE Methods of treatment using anti-erbB antibody-maytansinoid  
conjugates  
JOURNAL Patent: WO 0100244-A 2 04-JAN-2001;  
Genentech, Inc. (US)

FEATURES  
SOURCE Location/Qualifiers  
1..3768

/organism="Homo sapiens"  
/mol\_type="genomic DNA"

BASE COUNT 758 a 1170 c 1121 g 719 t  
ORIGIN

/db\_xref="taxon:9606"

Alignment Scores: 0.0683 Length: 3768  
Pred. No.: 83.00 Matches: 16  
Score: 83.00  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 6 Gaps: 0

JUNC\_SEQ8\_SEQ5 (1-30) x AX060704 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 2959 GTGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 3012

RESULT 10  
AX201817  
LOCUS AX201817 3768 bp DNA linear PAT 30-AUG-2001  
DEFINITION Sequence 1 from Patent WO0153463.  
ACCESSION AX201817  
VERSION AX201817.1 GI:15391666

KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS Cheever, M.A. and Hand-Zimmermann, S.  
TITLE Compounds and methods for prevention and treatment of her-2/ neu  
associated malignancies  
JOURNAL Patent: WO 0153463-A 1 26-JUL-2001;  
CORIAX CORPORATION (US)

FEATURES  
SOURCE Location/Qualifiers  
1..3768

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

1..3768

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC60126.1"

/db\_xref="GI:15391667"

/translation="MELAALCRWGLLLALLPGAASTOYCTGTDMKLRPLASPETHLD

MRLHYGQGVQGNLELYLPNALSIFLDIQEVOGVYVIAHNVQVQVPLORLIV  
RGTLFEDNVALAVLDNGDPLNNTPTVTPGASGGLRELQRLSLTEILKGLVILQRPQ  
LCYQDTILMKDIFHKNNQLALTLDITNRSRACHPCSPMCKGSRCSSESSEDQSLRT  
VCAGGCRKCGPLPTDCHEQCAAGCTGPKHSDCLACLFHNHSGICELCHCPALVTYT  
DTFESMPNPEGRYTFGACSVTACPNYVLTSDVGSCTLVCPLNHOEVTAEQGTORCEK  
SKPCARCYGLGMEHLREVRVTSANIQEFAGCKTFGSLAFPLPESFDGDPASNTAPL  
QPEQLVFTLEITGYLYISAWPDSLPLSVFQNLQVIRGRILHNGAYSLTLQGLGI  
SWIGLSRLRELGSGLAIHHNTHCFVHTVPDQLFNPQALLTANRPECEVGE  
LACHQICRHCWCPGPTQCVNCSQFLRQCEVCEVRLQSLPREYVNAHCLPCHPE  
CQPNQSVTCFGEADQCVACAHYKDPFCVRCPSGVKPDLSYIWKPFDEGACQ  
PCPINCTHSCVDLDDKCPAEORASPLTSLISAVGILLVVLGVVFGILKRRQKI  
RKYMRRLIQTETVELPPLPSGAMPNAQMIILKTELKRVKVLGSGAGTVYKGI  
PUGENVKIPVAKVLRNTSPKANKEILDEAYVMAGVGSFVSRILGICLTSTVQLVT  
QLMPYGCILLDHVRENRLGSLDLMWQIAKMSYLEDVRLVHLHLDNVLVKS  
NHVITDFGLRLDIDETEHADGKVPFKWMALESILRRFTHSDWSYGVTVME  
LMTFGAKPDVGPAREIPDLLEKRLPQPICTIDVYIMVWKWIDSCRRFRREL  
VSFESRMARDPQRFVIONEDLGPSPLDSTFYRSLLEDDMDGLDAEYLVPOQGF  
FCFDPAPGAGGVHHRSSSTRSGGDLTLGLEPSEEAAPRSLAPSGAGSDVDEG  
DUGMAAGKLQSLTHDPSFLQRYSEDPTVLPSTFDGIVFAPLTSPQPYVNPQDVR  
POPPSPREGPLPAARPAATLERPKTLSPKNGVVKDVFAGFVAVENPEYLTPOGGAA  
POPHPPPAFSPAFDNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGILDPV"

BASE COUNT 759 a 1171 c 1119 g 719 t  
ORIGIN

Alignment Scores:

Pred. No.: 0.0683 Length: 3768  
Score: 83.00 Matches: 16





Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Kreutzer, R., Limmer, S., Rost, S. and Hadwiger, P.  
Method for inhibiting the expression of a target gene  
Patent: WO 02055693-A 52 18-JUL-2002;  
Ribopharma AG (DE)

FEATURES

source  
1..3768  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 758 a 1170 c 1121 g 719 t  
ORIGIN

Alignment Scores:

Pred. No.:	0.0683	Length:	3768
Score:	83.00	Matches:	16
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	2
Query Match:	51.88%	Indels:	0
DB:	6	Gaps:	0

JUNC\_SEQ8\_SEQ5 (1-30) x AX481438 (1-3768)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 2959 GTGGTCATCCAGAAATGAGGACTTGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3012

Search completed: October 16, 2003, 11:01:50  
Job time : 436.208 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 234.794 Seconds  
(without alignments)  
3105.423 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ4  
Perfect score: 160  
Sequence: 1 ERGCPAEQRASPVTFQNEGLGPASPLDSTF 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV-xlp  
-Q/cgn2\_1/USPTO\_spool\_p/HOLLERAN480/runat\_15102003\_131913\_20548/app\_query.fasta\_1.4685  
-DB-EST -QFMT-fastp -SUFFIX-rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cgi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER-HOLLERAN480 -CGN\_1\_13645 -runat\_15102003\_131913\_20548 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

- Database :
- 1: em\_estba.\*
  - 2: em\_esthum.\*
  - 3: em\_estin.\*
  - 4: em\_estmu.\*
  - 5: em\_estov.\*
  - 6: em\_estpl.\*
  - 7: em\_estro.\*
  - 8: em\_htc.\*
  - 9: gb\_est1.\*
  - 10: gb\_est2.\*
  - 11: gb\_htc.\*
  - 12: gb\_est3.\*
  - 13: gb\_est4.\*
  - 14: gb\_est5.\*
  - 15: em\_estfun.\*
  - 16: em\_estom.\*
  - 17: em\_gss\_hum.\*
  - 18: em\_gss\_inv.\*
  - 19: em\_gss\_pln.\*
  - 20: em\_gss\_vrt.\*
  - 21: em\_gss\_fun.\*
  - 22: em\_gss\_nam.\*
  - 23: em\_gss\_mus.\*
  - 24: em\_gss\_pro.\*
  - 25: em\_gss\_rod.\*
  - 26: em\_gss\_phg.\*
  - 27: em\_gss\_vrl.\*
  - 28: gb\_gssl.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	83	51.9	278	10	BF923905	BF923905 QV4 -NT025
2	83	51.9	294	10	BF757857	BF757857 CM4 -CT057
c	83	51.9	337	10	BE091682	BE091682 IL2 -BT073
	83	51.9	348	9	A1909847	A1909847 QV -BT225-
5	83	51.9	378	14	CB266492	CB266492 1005398 H
6	83	51.9	423	12	BM703962	BM703962 UI -E-CK1-
7	83	51.9	464	2	HSM067221	Bx477782 Homo sapi
8	83	51.9	496	12	BM795191	BM795191 K-EST0076
9	83	51.9	505	12	BM854421	BM854421 K-EST0136
10	83	51.9	531	12	BM830170	BM830170 K-EST0103
11	83	51.9	536	14	CB129403	CB129403 K-EST0179
12	83	51.9	547	12	BM787824	BM787824 K-EST0066
13	83	51.9	567	2	HSM077766	Bx487579 Homo sapi
14	83	51.9	574	12	BM829991	BM829991 K-EST0103
15	83	51.9	630	2	HSM073298	Bx483130 Homo sapi
16	83	51.9	691	14	CB853376	CB853376 UI-CF-FNO
c	83	51.9	767	10	BF240297	BF240297 601905830
	83	51.9	849	14	CA489799	CA489799 AGENCOURT
19	83	51.9	852	14	CD516283	CD516283 AGENCOURT
20	83	51.9	866	14	CA488343	CA488343 AGENCOURT
21	83	51.9	893	14	CA455141	CA455141 AGENCOURT
22	83	51.9	894	14	CA454131	CA454131 AGENCOURT
23	83	51.9	899	14	CD558386	CD558386 AGENCOURT
24	83	51.9	916	13	B0845369	B0845369 AGENCOURT
25	83	51.9	919	13	B0845391	B0845391 AGENCOURT
26	83	51.9	924	14	CA489100	CA489100 AGENCOURT
27	83	51.9	932	14	CA487981	CA487981 AGENCOURT
28	83	51.9	984	14	CA489084	CA489084 AGENCOURT
29	83	51.9	1004	14	CD515356	CD515356 AGENCOURT
30	83	51.9	4715	11	AF318349	AF318349 Homo sapi
31	81	50.6	220	9	AW763060	AW763060 ur60c11.y
32	81	50.6	377	9	AW415355	AW415355 49593 MAR
33	81	50.6	483	4	BX530001	BX530001 RZPD Mus
34	81	50.6	518	10	BF039952	BF039952 Bp250015A
35	81	50.6	552	10	BF022141	BF022141 uy47c11.y
36	81	50.6	654	12	B1648646	B1648646 603277149
37	81	50.6	655	12	B1557778	B1557778 603236957
38	81	50.6	695	12	B1557170	B1557170 603239144
39	81	50.6	718	12	B1108229	B1108229 602902263
40	81	50.6	749	12	B1554781	B1554781 603236758
41	81	50.6	754	12	B1651822	B1651822 603299821
42	81	50.6	771	12	B1154609	B1154609 602904778
43	81	50.6	891	10	BF101641	BF101641 601753564
44	81	50.6	3110	11	AK031542	AK031542 Mus muscu
45	81	50.6	3372	11	BC046553	BC046553 Mus muscu

ALIGNMENTS

RESULT 1  
BF923905  
LOCUS QV4-NT0251-251100-599-all NF0251 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BF923905  
ACCESSION BF923905  
VERSION BF923905.1 GI:12319793  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 278)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
MEDLINE  
PUBMED  
10737800  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV4&t2=QV4-NT0251-251100-599-all&t3=2000-11-25&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 10  
High quality sequence stop: 278.  
Location/Qualifiers  
1. .278  
/organism="Homo sapiens"  
/mol\_type="mrna"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="NT0251"  
/note="Organ: nervous\_tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

#### FEATURES

source

BASE COUNT 52 a 78 c 79 g 69 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.0896 Length: 278  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x BF923905 (1-278)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 135 GTGGTCATCCAGATGAGGACTTGGGCCCGACGACCTTCCTGGACAGACCTTC 188

RESULT 2  
BF757857  
LOCUS CM4-CT0574-101100-428-f11 CT0574 Homo sapiens cDNA, mRNA sequence. EST 12-JAN-2001  
ACCESSION BF757857  
VERSION BF757857.1 GI:12105757  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
1 (bases 1 to 294)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
MEDLINE  
PUBMED  
10737800  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4&t2=CM4-CT0574-101100-428-f11&t3=2000-11-10&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 9  
High quality sequence stop: 294.  
Location/Qualifiers  
1. .294  
/organism="Homo sapiens"  
/mol\_type="mrna"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="CT0574"  
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

#### FEATURES

source

BASE COUNT 55 a 85 c 86 g 68 t  
ORIGIN

#### Alignment Scores:

Pred. No.: 0.0963 Length: 294  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x BF757857 (1-294)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 129 GTGGTCATCCAGATGAGGACTTGGGCCCGACGACCTTCCTGGACAGACCTTC 182

#### RESULT 3

BE091682/c

LOCUS IL2-BT0731-240400-070-E05 BT0731 Homo sapiens cDNA, mRNA sequence. EST 12-JUN-2000

ACCESSION BE091682

VERSION BE091682.1 GI:8482134

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS

1 (bases 1 to 337)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE  
PUBMED  
COMMENT

20202663  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=ft2=IL2-BT0731-240  
400-070-E05ft3=2000-04-24&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 336.

FEATURES  
source

1. .337  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="BT0731"

/note="Organ: breast; Vector: puc18; Site\_1: SmaI; Site\_2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No. 196  
716 - Ludwig Institute for Cancer Research) profiles  
into the pUC 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

BASE COUNT 72 a 101 c 96 g 68 t  
ORIGIN

Alignment Scores:

Pred. No.: 0.115 Length: 337  
Score: 83.00 Matches: 16  
Percent Similarity: 88.8% Conservative: 0  
Best Local Similarity: 88.8% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x BE091682 (1-337)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 148 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 95

RESULT 4

AI909847  
LOCUS QV-BT225-050599-038 BT225 Homo sapiens cDNA, mRNA linear EST 30-MAR-2000  
DEFINITION  
ACCESSION AI909847  
VERSION AI909847.1 GI:6500527  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
PUBMED 10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research

COMMENT

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/seq/gethtml.pl?tl=QVft2=QV-BT225-038.html  
ft3=050599ft4=1)

Seq primer: puc 18 forward.

FEATURES  
source

1. .348  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/sex="female"  
/dev\_stage="BT225"  
/clone\_lib="BT225"

/note="Organ: breast; Vector: puc18; Site\_1: SmaI; Site\_2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No. 196  
716 - Ludwig Institute for Cancer Research) profiles  
into the pUC 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

BASE COUNT 68 a 96 c 114 g 70 t  
ORIGIN

Alignment Scores:

Pred. No.: 0.12 Length: 348  
Score: 83.00 Matches: 16  
Percent Similarity: 88.8% Conservative: 0  
Best Local Similarity: 88.8% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 9 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x AI909847 (1-348)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
||| ||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 15 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 68

RESULT 5

CB266492 378 bp mRNA linear EST 20-FEB-2003  
LOCUS 1003398 Human Fat Cell 5'-Stretch plus cDNA Library Homo sapiens  
DEFINITION  
ACCESSION CB266492  
VERSION CB266492.1 GI:28441078  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Yang,R.-Z., Shuldiner,A. and Gong,D.-W.

EST analysis of human adipose gene expression  
Unpublished  
Contact: Gong Da-Wei  
Division of Endocrinology, Diabetes and Nutrition  
University of Maryland  
660 Redwood St. HH497, Baltimore, MD 21201, USA  
Tel: 410 706 1672  
Fax: 410 706 1622  
Email: dgong@medicine.umaryland.edu  
PCR Primers  
FORWARD: CTCGGGAAGCGCCATTGCTGTGTGT  
BACKWARD: AATAGCACTACTATAGGCGCAATTGG

Seq primer: GTGGTACCCGGGAATTC.  
Location/Qualifiers  
1. .378  
/organism="Homo sapiens"  
/mol\_type="mRNA"

FEATURES

source

```

/db_xref="taxon:9606"
/sex="Male and Female"
/tissue_type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
/notes="vector: lambdaTriplex"
BASE COUNT 71 a 110 c 117 g 80 t
ORIGIN
Alignment Scores:
Pred. No.: 0.133 Length: 378
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
Gaps: 0
JUNC_SEQ8_SEQ4 (1-30) x CB266492 (1-378)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 124 GTGTCATCCAGATGAGGACTTGGGCCCGCCAGCCAGTCCCTGGACAGCACCTTC 177
RESULT 6
LOCUS BM703962 423 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-CK1-afk-o-11-0-UI.r1 UI-E-CK1 Homo sapiens cDNA clone
VERSION BM703962
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 423)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel.: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1. .423
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CK1-afk-o-11-0-UI"
/tissue_type="Retina Foveal and Macular"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CK1"
/notes="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site.1: EcoR I; Site.2: Not I;
UI-E-CK1 is a normalized cDNA library containing the
following tissue(s): Retina Foveal and Macular. The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated

```

```

to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dF)18 tail. The sequence tag for this
library is GTCC. This library was created for the program,
Gene Discovery in the Visual System, supported by National
Eye Institute (NEI).
BASE COUNT 78 a 129 c 131 g 85 t
ORIGIN
Alignment Scores:
Pred. No.: 0.154 Length: 423
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
Gaps: 0
JUNC_SEQ8_SEQ4 (1-30) x BM703962 (1-423)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 147 GTGTCATCCAGATGAGGACTTGGGCCCGCCAGCCAGTCCCTGGACAGCACCTTC 200
RESULT 7
LOCUS HSM067221
DEFINITION HSM067221 standard; RNA; EST; 464 BP.
VERSION HSM067221
KEYWORDS AC BX477782;
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 464)
Bloeker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
Han M., Wiemann S.;
Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.
MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
This is the 5' sequence of the clone insert
CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
CC sequenced by GBF (National Research Centre for Biotechnology
CC Ltd., Braunschweig/Germany) within the cDNA sequencing
CC consortium of the German Genome Project.
CC No sl sequence available.
CC This clone (DKFZp686M12198) is available at the RZPD in Berlin.
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,
CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX
Key Location/Qualifiers
FH source 1. .464
FT /db_xref="taxon:9606"
FT /mol_type="mRNA"
FT /organism="Homo sapiens"
FT /clone="DKFZp686M12198"
FT /clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
FT DH10B; sites SfiIA + SfiIB"
FT /dev_stage="adult"

```



## ORIGIN

Alignment Scores: 0.193 Length: 505  
 Pred. No.: 83.00 Matches: 16  
 Score: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 51.88% Indels: 0  
 DB: 12 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x BM854421 (1-505)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 DB 334 GTGGTCATCCAGAAATGAGGACTTGGCCAGCCAGTCCCTTGGACAGCACCTTC 387

## RESULT 10

BM830170  
 LOCUS K-EST0103592 S21SNU520s1 531 bp mRNA linear EST 06-MAR-2002  
 DEFINITION 5', mRNA sequence.

ACCESSION BM830170.1 GI:19186579

## KEYWORDS

SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 531)

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and

Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 19 row: D column: 03

High quality sequence stop: 531.

## FEATURES

source

1..531

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S21SNU520s1-19-D03"

/sex="F"

/tissue\_type="Stomach"

/cell\_type="Floating aggregates"

/cell\_line="SNU-520"

/lab\_host="Top10F"

/clone\_lib="S21SNU520s1"

/note="Organ: Stomach; Vector: pTZ18RP1; Site.1: EcoRI;

Site.2: NotI; The poly (A)+ RNA was dephosphorylated with

bacterial alkaline phosphatase (BAP) and then decapped

with tobacco acid pyrophosphatase (TAP). The decapped

intact mRNA was ligated with DNA-RNA linker including EcoR

I site by treatment of T4 RNA ligase and the first strand

cDNA was synthesized from oligo dt-selected mRNA by

priming with dt-tailed vector. The dt-tailed vector was

adjusted to have about 60nt. The cDNA vector was

circularized with E. coli DNA ligase after digestion of

EcoRI which site is also included in vector. An RNA strand

converted to a DNA strand by Okayama-Berg method. The

obtained cDNA vectors were used for transformation of

competent cells E. coli Top10F' by electroporation method.

The cDNA libraries constructed by this method are

full-length enriched cDNA library. After analyzing and

sequencing about 2,000 ~ 3,000 colonies in original cDNA

library, the abundant cDNAs were selected and amplified by

PCR reaction using vector region primer including T7  
 promoter as 5' primer and N(dt)14 as 3' primer. The PCR  
 products were used as template for synthesis of  
 biotinylated single stranded RNA by in vitro transcription  
 reaction. The synthesized cDNA probes were hybridized with  
 antisense single stranded cDNAs prepared from original  
 library and incubated with avidin-gel. After removing  
 DNA-RNA hybrids by centrifuge, the subtracted cDNA  
 libraries were constructed by transfection of the  
 remaining DNA into competent cells E. coli Top10F' with  
 electroporation method."

BASE COUNT 102 a 149 c 166 g 114 t

## ORIGIN

Alignment Scores: 0.206 Length: 531  
 Pred. No.: 83.00 Matches: 16  
 Score: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 51.88% Indels: 0  
 DB: 12 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x BM830170 (1-531)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 DB 334 GTGGTCATCCAGAAATGAGGACTTGGCCAGCCAGTCCCTTGGACAGCACCTTC 387

## RESULT 11

CB129403

LOCUS K-EST0179037 C1SNU17 Homo sapiens cDNA clone C1SNU17-5-C08 5', mRNA

DEFINITION

sequence.

ACCESSION CB129403

VERSION CB129403.1 GI:28093000

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 536)

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and

Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 5 row: C column: 08

High quality sequence stop: 536.

## FEATURES

source

1..536

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="C1SNU17-5-C08"

/sex="F"

/tissue\_type="uterine"

/cell\_type="Epithelial"

/cell\_line="SNU-17"

/lab\_host="Top10F"

/clone\_lib="C1SNU17"

/note="Organ: Cervix; Vector: pCNS-D2; Site.1: EcoRI;

Site.2: NotI; The poly (A)+ RNA was dephosphorylated with

bacterial alkaline phosphatase (BAP) and then decapped

with tobacco acid pyrophosphatase (TAP). The decapped

intact mRNA was ligated with DNA-RNA linker including

EcoRI site by treatment of T4 RNA ligase and the first

strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 105 a 150 c 167 g 114 t

ORIGIN

#### Alignment Scores:

Pred. No.: 0.209 Length: 536  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 14 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x CB129403 (1-536)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 339 GTGGTCATCCAGAAATGAGGACTTGGCCCGAGCAGCTTCCTTGGACAGCACCTTC 392

#### RESULT 12

BM787824

LOCUS K-EST0066898 S11SNUI Homo sapiens cDNA clone S11SNUI-23-G01 5',  
DEFINITION mRNA sequence.

ACCESSION BM787824.1 GI:19136056

VERSION EST.

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 547)

REFERENCE Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
AUTHORS Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 23 row: G column: 01

High quality sequence stop: 547.

Location/Qualifiers

1..547

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S11SNUI-23-G01"

/sex="M"

/tissue\_type="Stomach"

/cell\_type="Lymphoblast-like"

/cell\_line="SNU-1"

/lab\_host="Top10F"

/clone\_lib="S11SNUI"

/note="Organ: Stomach; Vector: pME18-FL3; Site:1: XhoI;

Site:2: XhoI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including SfiI  
site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized with Superscript II using SfiI

oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into DraIII- digested pME18S-FL3 vector. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 102 a 187 c 163 g 95 t

ORIGIN

#### Alignment Scores:

Pred. No.: 0.214 Length: 547  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x BM787824 (1-547)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 39 GTGGTCATCCAGAAATGAGGACTTGGCCCGAGCAGCTTCCTTGGACAGCACCTTC 92

#### RESULT 13

HSN077766

ID HSN077766 standard; RNA; EST: 567 BP.

XX AC BX487579;

XX SV BX487579.1

XX XX

DT 09-MAY-2003 (Rel. 75, Created)

DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)

DE Homo sapiens mRNA; EST DKFZp686M16257\_r1 (from clone DKFZp686M16257)  
KW EST; expressed sequence tag.

XX Homo sapiens (human)

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

XX [1]

RP 1-567

RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,

RT Han M., Wiemann S.;

RL Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.

RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY

XX This is the 5' sequence of the clone insert

CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

CC sequenced by BMFZ (Biomedical Research Center at the

CC Heinrich-Heine-University, Duesseeldorf/Germany) within the cDNA

CC sequencing consortium of the German genome Project.

CC No si sequence available.

CC This clone (DKFZp686M16257) is available at the RZPD in Berlin.

CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,

CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

XX Key

Location/Qualifiers

1..567

/db\_xref="taxon:9606"

/mol\_type="mRNA"

/organism="Homo sapiens"

/clone="DKFZp686M16257"

/clone\_lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host

DL10B; sites SfiIA + SfiIB"

/dev\_stage="adult"

FT source

FT /tissue\_type="cDNA-collection"

XX Sequence 567 BP; 115 A; 146 C; 184 G; 122 T; 0 other;

#### Alignment Scores:

Pred. No.: 0.224 Length: 567  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 2 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x HSM077766 (1-567)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrphe 30

Db 468 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTGGACACGACCTTC 521

#### RESULT 14

BM829991

LOCUS 574 bp mRNA linear EST 06-MAR-2002  
DEFINITION K-EST0103390 S21SNU520s1 Homo sapiens cDNA clone S21SNU520s1-17-C05  
5', mRNA sequence.

ACCESSION BM829991

VERSION BM829991.1 GI:19186400

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 574)

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished

COMMENT Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 17 row: C column: 05

High quality sequence stop: 574.

#### FEATURES

source

1..574

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S21SNU520s1-17-C05"

/sex="F"

/tissue\_type="Stomach"

/cell\_type="Floating aggregates"

/cell\_line="SNU-520"

/lab\_host="Top10F"

/clone\_lib="S21SNU520s1"

/note="Organ: Stomach; Vector: pTZ19Rpl; Site:1: EcoRI;

Site:2: NotI; The poly (A)+ RNA was dephosphorylated with

bacterial alkaline phosphatase (BAP) and then decapped

with tabacco acid pyrophosphatase (TAP). The decapped

intact mRNA was ligated with DNA-RNA linker including EcoR

I site by treatment of T4 RNA ligase and the first strand

cDNA was synthesized from oligo dt-selected mRNA by

priming with dt-tailed vector. The dt-tailed vector was

adjusted to have about 60nt. The cDNA vector was

circularized with E. coli DNA ligase after digestion of

EcoRI which site is also included in vector. An RNA strand

converted to a DNA strand by Okayama-Berg method. The

converted cDNA vectors were used for transformation of

competent cells E. coli Top10F' by electroporation method.

The cDNA libraries constructed by this method are

full-length enriched cDNA library. After analyzing and

sequencing about 2,000 ~ 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 PCR promotor as 5' primer and N(dT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F' with electroporation method."

BASE COUNT 109 a 159 c 185 g 121 t

ORIGIN

#### Alignment Scores:

Pred. No.: 0.228 Length: 574  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x BM829991 (1-574)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrphe 30

Db 334 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTGGACACGACCTTC 387

#### RESULT 15

HSM073298

ID HSM073298 standard; RNA; EST: 630 BP.

XX BX483130;

XX BX483130.1

SV

XX

XX

DT 09-MAY-2003 (Rel. 75, Created)

DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)

XX

DE Homo sapiens mRNA; EST DKFZp686N09234\_r1 (from clone DKFZp686N09234)

XX

XX EST; expressed sequence tag.

XX

OS Homo sapiens (human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

OC Eutheria; Primates; Catarrhini; Homnidae; Homo.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX



FT /clone="DKFZp686N09234"  
FT /clone\_lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host  
FT DH10B; sites SfiIA + SfiIB"  
FT /dev\_stage="adult"  
FT /tissue\_type="cDNA-collection"  
XX

SQ Sequence 630 BP; 124 A; 174 C; 196 G; 133 T; 3 other;

Alignment Scores:  
Pred. No.: 0.257 Length: 630  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 2 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x HSM073298 (1-630)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
||| |||||  
Db 439 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCTTGGACAGCACCTTC 492

Search completed: October 16, 2003, 17:04:06  
Job time : 236.794 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 16, 2003, 11:08:07 ; Search time 31.0111 Seconds  
(without alignments)  
2540.503 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ4

Perfect score: 160

Sequence: 1 ERGCPAEQRAQSPVTFQNEGLGPASPLDSTF 30

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO\_spool\_p/HOLLERAN480/runat\_15102003\_131915\_20662/app\_query.fasta\_1.4685  
-DB=PublishedApplications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-NAXLEN=2000000000 -USER=HOLLERAN480 -CGN\_1\_1\_397 -runat\_15102003\_131915\_20662  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database : PublishedApplications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	83	51.9	1115	14	US-10-102-806-165

2	83	51.9	1713	12	US-10-378-393-14	Sequence 14, Appl
3	83	51.9	1755	10	US-09-930-125-6	Sequence 6, Appl
4	83	51.9	1767	10	US-09-930-125-4	Sequence 4, Appl
5	83	51.9	1773	10	US-09-930-125-7	Sequence 7, Appl
6	83	51.9	1806	10	US-09-930-125-5	Sequence 5, Appl
7	83	51.9	2411	12	US-10-378-393-10	Sequence 10, Appl
8	83	51.9	3765	12	US-10-207-498-5	Sequence 5, Appl
9	83	51.9	3768	9	US-09-811-123-8	Sequence 8, Appl
10	83	51.9	3768	9	US-09-811-115-2	Sequence 2, Appl
11	83	51.9	3768	10	US-09-854-356-9	Sequence 9, Appl
12	83	51.9	3768	10	US-09-930-125-1	Sequence 1, Appl
13	83	51.9	3768	12	US-10-313-644-1	Sequence 1, Appl
14	83	51.9	4473	11	US-09-441-411-5	Sequence 5, Appl
15	83	51.9	4473	12	US-10-101-510-81	Sequence 81, Appl
16	83	51.9	4473	14	US-10-146-473-32	Sequence 32, Appl
17	83	51.9	4473	14	US-10-207-655-44	Sequence 44, Appl
18	83	51.9	4530	10	US-09-877-177-11	Sequence 11, Appl
19	83	51.9	4530	12	US-10-007-926A-119	Sequence 119, App
20	83	51.9	4530	12	US-10-101-510-124	Sequence 124, App
21	83	51.9	4530	12	US-10-338-730-1	Sequence 1, Appl
22	83	51.9	4530	14	US-10-177-293-125	Sequence 125, App
23	83	51.9	4543	10	US-09-769-508-1	Sequence 1, Appl
24	83	51.9	4606	12	US-09-971-392-70	Sequence 70, Appl
25	83	51.9	4642	14	US-10-198-846-10896	Sequence 10896, A
26	83	51.9	9274	9	US-09-811-123-7	Sequence 7, Appl
27	83	51.9	9274	9	US-09-811-115-1	Sequence 1, Appl
28	81	50.6	3771	10	US-09-854-356-11	Sequence 11, Appl
29	81	50.6	3955	10	US-09-870-759-117	Sequence 117, App
30	81	50.6	3955	10	US-09-854-356-10	Sequence 10, Appl
31	81	50.6	3955	12	US-09-751-708A-117	Sequence 117, App
32	66	41.2	201	13	US-10-109-213-3	Sequence 3, Appl
33	58.5	36.6	143601	12	US-09-855-824-3	Sequence 3, Appl
34	58	36.2	1938	12	US-10-422-264-25	Sequence 25, Appl
35	58	36.2	2316	12	US-10-422-264-27	Sequence 27, Appl
36	58	36.2	2604	12	US-10-422-264-23	Sequence 23, Appl
37	58	36.2	3320	10	US-09-862-658-1	Sequence 1, Appl
38	58	36.2	3320	14	US-10-175-696-22	Sequence 22, Appl
39	58	36.2	3384	12	US-10-422-264-29	Sequence 29, Appl
40	57.5	35.9	363	11	US-09-918-995-28706	Sequence 28706, A
c 41	57	35.6	23432	9	US-09-764-869-1332	Sequence 1332, Ap
c 42	57	35.6	23432	14	US-10-091-504-1332	Sequence 1332, Ap
c 43	56.5	35.3	1557	10	US-09-774-434-6	Sequence 6, Appl
c 44	56.5	35.3	22452	10	US-09-764-868-1487	Sequence 1487, Ap
c 45	56.5	35.3	22452	10	US-09-764-868-1489	Sequence 1489, Ap

#### ALIGNMENTS

RESULT 1  
US-10-102-806-165  
; Sequence 165, Application US/10102806  
; Publication No. US2003005442A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA103P1C1  
; CURRENT APPLICATION NUMBER: US/10/102,806  
; CURRENT FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 09/925,298  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05881  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 846  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 165  
; LENGTH: 1115  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (390)

; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (394)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-102-806-165

Alignment Scores:  
Pred. No.: 0.00141 Length: 1115  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 14 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x US-10-102-806-165 (1-1115)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 446 GTGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 499

## RESULT 2

US-10-378-393-14  
; Sequence 14, Application US/10378393  
; Publication No. US20030182668A1  
; GENERAL INFORMATION:  
; APPLICANT: Bol, David K.  
; APPLICANT: Carboni, Joan M.  
; APPLICANT: Rowley, Ronald B.  
; APPLICANT: Wong, Tai W.  
; APPLICANT: Lee, Francis Y.  
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVATED  
; FILE REFERENCE: D0254 NP  
; CURRENT APPLICATION NUMBER: US/10/378,393  
; CURRENT FILING DATE: 2003-03-03  
; PRIOR APPLICATION NUMBER: US 60/360,889  
; PRIOR FILING DATE: 2002-03-01  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 14  
; LENGTH: 1713  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-378-393-14

Alignment Scores:  
Pred. No.: 0.00225 Length: 1713  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x US-10-378-393-14 (1-1713)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 904 GTGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 957

## RESULT 3

US-09-930-125-6  
; Sequence 6, Application US/09930125  
; Publication No. US20020193329A1  
; GENERAL INFORMATION:  
; APPLICANT: Hand-Zimmerman, Susan  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Vedvick, Thomas S.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS  
; OF HER-2/NEU-ASSOCIATED MALIGNANCIES

; FILE REFERENCE: 210121.544  
; CURRENT APPLICATION NUMBER: US/09/930.125  
; CURRENT FILING DATE: 2001-08-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 1755  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-930-125-6

Alignment Scores:  
Pred. No.: 0.00231 Length: 1755  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x US-09-930-125-6 (1-1755)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 937 GTGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 990

## RESULT 4

US-09-930-125-4  
; Sequence 4, Application US/09930125  
; Publication No. US20020193329A1  
; GENERAL INFORMATION:  
; APPLICANT: Hand-Zimmerman, Susan  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Vedvick, Thomas S.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS  
; OF HER-2/NEU-ASSOCIATED MALIGNANCIES  
; FILE REFERENCE: 210121.544  
; CURRENT APPLICATION NUMBER: US/09/930.125  
; CURRENT FILING DATE: 2001-08-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 1767  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-930-125-4

Alignment Scores:  
Pred. No.: 0.00233 Length: 1767  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x US-09-930-125-4 (1-1767)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 937 GTGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 990

## RESULT 5

US-09-930-125-7  
; Sequence 7, Application US/09930125  
; Publication No. US20020193329A1  
; GENERAL INFORMATION:  
; APPLICANT: Hand-Zimmerman, Susan  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Lodes, Michael J.

```
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-7

Alignment Scores:
Pred. No.: 0.00233 Length: 1773
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-930-125-7 (1-1773)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 961 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 1014

RESULT 6
US-09-930-125-5
; Sequence 5, Application US/099301025
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1806
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-5

Alignment Scores:
Pred. No.: 0.00238 Length: 1806
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-930-125-5 (1-1806)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 994 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 1047

RESULT 7
US-10-378-393-10
; Sequence 10, Application US/10378393
; Publication No. US20030182668A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVAT
; FILE REFERENCE: D0254 NP
; CURRENT APPLICATION NUMBER: US/10/378,393
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,889
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-378-393-10

Alignment Scores:
Pred. No.: 0.00327 Length: 2411
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-10-378-393-10 (1-2411)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 1602 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 1655

RESULT 8
US-10-207-498-5
; Sequence 5, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; FILE REFERENCE: 30448.103-US-U1
; CURRENT APPLICATION NUMBER: US/10/207,498
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-207-498-5

Alignment Scores:
Pred. No.: 0.00532 Length: 3765
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-10-207-498-5 (1-3765)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
```

```

Db      2959 GTGTCATCCAGAATGAGGACTTGGGCCCCAGCCAGTCCCTTGGACAGCACCTTC 3012
|||||
RESULT 9
US-09-811-123-8
; Sequence 8, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
; TITLE OF INVENTION: ANTIBODY-MAYTANSINOID CONJUGATES
; FILE REFERENCE: GENENT.073A2
; CURRENT APPLICATION NUMBER: US/09/811.123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-123-8

Alignment Scores:
Pred. No.:      0.00532      Length:      3768
Score:          83.00      Matches:      16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:    51.88%      Indels:      0
DB:             9      Gaps:      0

JUNC_SEQ8_SEQ4 (1-30) x US-09-811-123-8 (1-3768)

Qy      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
|||||
Db      2959 GTGTCATCCAGAATGAGGACTTGGGCCCCAGCCAGTCCCTTGGACAGCACCTTC 3012
|||||
RESULT 10
US-09-811-115-2
; Sequence 2, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811.115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-115-2

Alignment Scores:
Pred. No.:      0.00532      Length:      3768
Score:          83.00      Matches:      16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:    51.88%      Indels:      0
DB:             9      Gaps:      0

JUNC_SEQ8_SEQ4 (1-30) x US-09-811-115-2 (1-3768)

Db      2959 GTGTCATCCAGAATGAGGACTTGGGCCCCAGCCAGTCCCTTGGACAGCACCTTC 3012
|||||
RESULT 11
US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

Alignment Scores:
Pred. No.:      0.00532      Length:      3768
Score:          83.00      Matches:      16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:    51.88%      Indels:      0
DB:             10      Gaps:      0

JUNC_SEQ8_SEQ4 (1-30) x US-09-854-356-9 (1-3768)

Qy      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
|||||
Db      2959 GTGTCATCCAGAATGAGGACTTGGGCCCCAGCCAGTCCCTTGGACAGCACCTTC 3012
|||||
RESULT 12
US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS  
 ; FILE REFERENCE: 210121.544  
 ; CURRENT APPLICATION NUMBER: US/09/930,125  
 ; CURRENT FILING DATE: 2001-08-14  
 ; NUMBER OF SEQ ID NOS: 25  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 3768  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)...(3765)  
 US-09-930-125-1

Alignment Scores:  
 Pred. No.: 0.00532 Length: 3768  
 Score: 83.00 Matches: 16  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 51.88% Indels: 0  
 DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x US-09-930-125-1 (1-3768)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 ||| |||||  
 Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 3012

RESULT 13

US-10-313-644-1  
 ; Sequence 1, Application US/10313644  
 ; Publication No. US20030157119A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gaiger, Alexander  
 ; APPLICANT: Cheever, Martin A.  
 ; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL  
 ; FILE REFERENCE: 210121.483C3  
 ; CURRENT APPLICATION NUMBER: US/10/313,644  
 ; CURRENT FILING DATE: 2002-12-04  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 3768  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)...(3765)  
 US-10-313-644-1

Alignment Scores:  
 Pred. No.: 0.00532 Length: 3768  
 Score: 83.00 Matches: 16  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 51.88% Indels: 0  
 DB: 12 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x US-10-313-644-1 (1-3768)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 ||| |||||  
 Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 3012

RESULT 14

US-09-441-411-5  
 ; Sequence 5, Application US/09441411  
 ; Publication No. US20030008342A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Scholler, Nathalie B.  
 ; APPLICANT: Disis, Mary L.  
 ; APPLICANT: Hellstrom, Ingegerd  
 ; APPLICANT: Hellstrom, Karl Erik  
 ; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES  
 ; FILE REFERENCE: 730033.409  
 ; CURRENT APPLICATION NUMBER: US/09/441,411  
 ; CURRENT FILING DATE: 1999-11-16  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5  
 ; LENGTH: 4473  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-441-411-5

Alignment Scores:  
 Pred. No.: 0.00642 Length: 4473  
 Score: 83.00 Matches: 16  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 51.88% Indels: 0  
 DB: 11 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x US-09-441-411-5 (1-4473)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 ||| |||||  
 Db 3133 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 3186

RESULT 15

US-10-101-510-81  
 ; Sequence 81, Application US/10101510  
 ; Publication No. US20030148295A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WANG, JACKSON  
 ; APPLICANT: WANG, YIXIN  
 ; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE  
 ; FILE REFERENCE: 15117.0012  
 ; CURRENT APPLICATION NUMBER: US/10/101,510  
 ; CURRENT FILING DATE: 2002-03-20  
 ; PRIOR APPLICATION NUMBER: 60/276,947  
 ; PRIOR FILING DATE: 2001-03-20  
 ; NUMBER OF SEQ ID NOS: 805  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 81  
 ; LENGTH: 4473  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-101-510-81

Alignment Scores:  
 Pred. No.: 0.00642 Length: 4473  
 Score: 83.00 Matches: 16  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 51.88% Indels: 0  
 DB: 12 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x US-10-101-510-81 (1-4473)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 ||| |||||  
 Db 3133 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 3186

Search completed: October 17, 2003, 03:53:11  
 Job time : 35.0111 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 15, 2003, 22:55:19 ; Search time 7.60488 Seconds

(without alignments)  
1741.185 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ4

Perfect score: 160

Sequence: 1 ERGCCPAERASPVTFQNEGLGASPLDSTF 30

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delpop 6.0, Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n model -DEV=xlp

-Q/cgn2\_1/USPTO.spool\_p/HOLLERAN480/runat\_15102003\_131913\_20570/app\_query.fasta\_1.4685

-DB=Issued\_Patents\_NA -OFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi

-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000

-USER=HOLLERAN480@cgn.1.1.403@runat\_15102003\_131913\_20570 -NCPU=6 -ICPU=3

-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA: \*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*

2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq: \*

3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq: \*

4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq: \*

5: /cgn2\_6/ptodata/2/ina/pctus\_COMB.seq: \*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	51.9	3768	2	US-08-625-101-1
2	83	51.9	3768	2	US-08-356-786-1
3	83	51.9	4473	2	US-09-048-804-1
4	83	51.9	4473	3	US-09-056-105-26
5	83	51.9	4530	1	US-08-229-515A-9
6	83	51.9	4530	1	US-08-645-865-9
7	83	51.9	4530	4	US-09-167-322-4
8	83	51.9	4530	4	US-09-527-487-1
9	83	51.9	4530	4	US-09-877-177A-11
10	81	50.6	3955	1	US-08-229-515A-14
11	81	50.6	3955	1	US-08-645-865-14
12	68	42.5	2385	2	US-09-146-283-3

13	68	42.5	2385	3	US-08-579-823A-3	Sequence 3, Appli
14	68	42.5	2385	3	US-09-344-195-3	Sequence 3, Appli
15	66	41.2	153	3	US-08-776-251-3	Sequence 3, Appli
16	66	41.2	201	4	US-09-200-355-3	Sequence 3, Appli
17	66	41.2	816	3	US-08-776-251-10	Sequence 10, Appl
18	58	36.2	1938	4	US-09-547-435-25	Sequence 25, Appl
19	58	36.2	2316	4	US-09-547-435-27	Sequence 27, Appl
20	58	36.2	2604	4	US-09-547-435-23	Sequence 23, Appl
21	58	36.2	3384	4	US-09-547-435-29	Sequence 29, Appl
22	56.5	35.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
23	56.5	35.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
24	56	35.0	1368	4	US-09-328-352-430	Sequence 430, App
25	55.5	34.7	2691	4	US-09-020-743-1	Sequence 1, Appli
26	54.5	34.1	536165	4	US-09-214-808-1	Sequence 1, Appli
27	54	33.8	42571	4	US-09-810-347-3	Sequence 3, Appli
28	53.5	33.4	634	1	US-08-451-947-1	Sequence 1, Appli
29	53.5	33.4	634	2	US-08-424-826A-1	Sequence 1, Appli
30	53.5	33.4	634	3	US-08-928-694-1	Sequence 1, Appli
31	53.5	33.4	634	4	US-08-450-842-1	Sequence 1, Appli
32	53.5	33.4	634	4	US-08-451-390-1	Sequence 1, Appli
33	53.5	33.4	634	5	PCT-US91-06950-1	Sequence 1, Appli
34	53.5	33.4	864	3	US-09-022-983-3	Sequence 3, Appli
35	53.5	33.4	1404	1	US-07-796-106-22	Sequence 22, Appl
36	53.5	33.4	2340	3	US-09-022-983-4	Sequence 4, Appli
37	53.5	33.4	2477	3	US-09-490-692-3	Sequence 3, Appli
38	53	33.1	960	3	US-08-651-136C-1	Sequence 1, Appli
39	53	33.1	960	4	US-09-229-911A-1	Sequence 1, Appli
40	53	33.1	5046	4	US-09-548-938A-5	Sequence 5, Appli
41	53	33.1	11282	4	US-09-754-250-3	Sequence 3, Appli
42	51.5	32.2	3225	4	US-09-358-856C-2	Sequence 2, Appli
43	51	31.9	4092	3	US-09-306-595C-5	Sequence 5, Appli
44	51	31.9	4092	4	US-09-925-388-5	Sequence 5, Appli
45	51	31.9	7672	4	US-09-220-132-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-08-625-101-1

; Sequence 1, Application US/08625101

; Patent No. 5869445

; GENERAL INFORMATION:

; APPLICANT: Cheever, Martin A.

; APPLICANT: Disis, Mary L.

; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE

; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION

; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu

; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/625,101

; FILING DATE: 01-APR-1996

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Sharkey, Richard G.

; REGISTRATION NUMBER: 32,629

; REFERENCE/DOCKET NUMBER: 920010.448C7

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 1:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3765
US-08-625-101-1

Alignment Scores:
Pred. No.: 0.00745 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-08-625-101-1 (1-3768)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3012

RESULT 2
US-08-356-786-1
; Sequence 1, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356.786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3768
; OTHER INFORMATION: /note= "product = "cerb-b2"
```

```
US-08-356-786-1

Alignment Scores:
Pred. No.: 0.00745 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-08-356-786-1 (1-3768)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3012

RESULT 3
US-09-048-804-1
; Sequence 1, Application US/09048804
; Patent No. 5968748
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
; TITLE OF INVENTION: HUMAN HER-2 EXPRESSION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748ris LLP
; STREET: One Liberty Place - 48th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/048.804
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2913
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4473 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; ANTI-SENSE: NO
US-09-048-804-1

Alignment Scores:
Pred. No.: 0.0092 Length: 4473
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-048-804-1 (1-4473)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3133 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3186

RESULT 4
US-09-056-105-26
```

; Sequence 26, Application US/09056105  
; Patent No. 6287569  
; GENERAL INFORMATION:  
; APPLICANT: KIPPS, THOMAS J.  
; APPLICANT: WU, YUNQI  
; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR  
; TITLE OF INVENTION: PROCESSING  
; FILE REFERENCE: 233/221  
; CURRENT APPLICATION NUMBER: US/09/056,105  
; CURRENT FILING DATE: 1998-04-06  
; EARLIER APPLICATION NUMBER: 60/043,467  
; EARLIER FILING DATE: 1997-04-10  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 26  
; LENGTH: 4473  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-056-105-26

Alignment Scores:  
Pred. No.: 0.0092 Length: 4473  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 3 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x US-09-056-105-26 (1-4473)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
|||  
Db 3133 GTGGTCATCCAGAAATGAGGACTTGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3186

## RESULT 5

US-08-229-515A-9  
; Sequence 9, Application US/08229515A  
; Patent No. 5518885  
; GENERAL INFORMATION:  
; APPLICANT: RAZIUDIN  
; APPLICANT: SARKAR, FAZLUL H  
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN  
; TITLE OF INVENTION: NEOPLASTIC DISEASE  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG PC  
; STREET: 127 Peachtree Street, Suite 1200  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: usa  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/229,515A  
; FILING DATE: 19 APR 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PERRYMAN, DAVID G  
; REGISTRATION NUMBER: 33,438  
; REFERENCE/DOCKET NUMBER: 1414.608  
; TELEPHONE: 404-688-9880  
; TELEFAX: 404-688-9880  
; INFORMATION FOR SEQ ID NO: 9:  
; LENGTH: 4530 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-229-515A-9

Alignment Scores:  
Pred. No.: 0.00935 Length: 4530  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 1 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x US-08-229-515A-9 (1-4530)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
|||  
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGCCCGCCAGTCCCTTGGACAGCACCTTC 3162

## RESULT 6

US-08-645-865-9  
; Sequence 9, Application US/08645865  
; Patent No. 5654406  
; GENERAL INFORMATION:  
; APPLICANT: RAZIUDIN  
; APPLICANT: SARKAR, FAZLUL H  
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN  
; TITLE OF INVENTION: NEOPLASTIC DISEASE  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG PC  
; STREET: 127 Peachtree Street, Suite 1200  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: usa  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/645,865  
; FILING DATE: 14 MAY 1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PERRYMAN, DAVID G  
; REGISTRATION NUMBER: 33,438  
; REFERENCE/DOCKET NUMBER: 1414.608  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404-688-0770  
; TELEFAX: 404-688-9880  
; INFORMATION FOR SEQ ID NO: 9:  
; LENGTH: 4530 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-645-865-9

Alignment Scores:  
Pred. No.: 0.00935 Length: 4530  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 1 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x US-08-645-865-9 (1-4530)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
|||  
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGCCCGCCAGTCCCTTGGACAGCACCTTC 3162

## RESULT 7

US-09-167-322-4

```
; Sequence 4, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-167-322-4

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-167-322-4 (1-4530)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACACGACCTTC 3162

RESULT 8
US-09-527-487-1
; Sequence 1, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(3915)
US-09-527-487-1

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-527-487-1 (1-4530)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACACGACCTTC 3162

RESULT 9
US-09-877-177A-11
; Sequence 11, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-11

Alignment Scores:
Pred. No.: 0.00935 Length: 4530
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-09-877-177A-11 (1-4530)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACACGACCTTC 3162

RESULT 10
US-08-229-515A-14
; Sequence 14, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

; TOPOLOGY: linear
US-08-645-865-14

Alignment Scores:
Pred. No.: 0.0158 Length: 3955
Score: 81.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.62% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x US-08-645-865-14 (1-3955)

Qy 1 GluArgGlyCysProAlaGluInArgAlaSerProValThrPhe 15
|||||
Db 1946 GAACGAGGCTGCCAGCAGCAGACAGACGACCGGTGACATTC 1990

RESULT 12
US-09-146-283-3
; Sequence 3, Application US/09146283
; Patent No. 5976546
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-09-146-283-3

Alignment Scores:
Pred. No.: 0.766 Length: 2385
Score: 68.00 Matches: 17
Percent Similarity: 66.67% Conservative: 3
Best Local Similarity: 56.67% Mismatches: 6
Query Match: 42.50% Indels: 4
DB: 2 Gaps: 2

JUNC_SEQ8_SEQ4 (1-30) x US-09-146-283-3 (1-2385)

```

```
Oy 1 GluArgGlyCysProAlaGluInArgAlaSerProValThrPheGlnAsnGluAspLeu 20
    ::::|||||
Db 1925 GACAAGGCTGCCCGCCGAGCAGAGCCAGCCCTCTGACGTCCCTCGAG----- 1975

Oy 21 GlyProAla---SerProLeuAspSerThr 29
    |||||
Db 1976 GCACCCGCGCTCGCCCGAGCCCGCAGCACA 2005

RESULT 13
US-08-579-823A-3
; Sequence 3, Application US/08579823A
; Patent No. 6080409
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; TITLE OF INVENTION: Immunostimulatory Composition and Method
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,823A
; FILING DATE: 03-DEC-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-08-579-823A-3

Alignment Scores:
Pred. No.: 0.766 Length: 2385
Score: 68.00 Matches: 17
Percent Similarity: 66.67% Conservative: 3
Best Local Similarity: 56.67% Mismatches: 6
Query Match: 42.50% Indels: 4
DB: 3 Gaps: 2

JUNC_SEQ8_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)

Oy 1 GluArgGlyCysProAlaGluInArgAlaSerProValThrPheGlnAsnGluAspLeu 20
    ::::|||||
Db 1925 GACAAGGCTGCCCGCCGAGCAGAGCCAGCCCTCTGACGTCCCTCGAG----- 1975

Oy 21 GlyProAla---SerProLeuAspSerThr 29
    |||||
Db 1976 GCACCCGCGCTCGCCCGAGCCCGCAGCACA 2005
```

```
RESULT 14
US-09-344-195-3
; Sequence 3, Application US/09344195
; Patent No. 6210662
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,195
; FILING DATE: 24-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-09-344-195-3

Alignment Scores:
Pred. No.: 0.766 Length: 2385
Score: 68.00 Matches: 17
Percent Similarity: 66.67% Conservative: 3
Best Local Similarity: 56.67% Mismatches: 6
Query Match: 42.50% Indels: 4
DB: 3 Gaps: 2

JUNC_SEQ8_SEQ4 (1-30) x US-09-344-195-3 (1-2385)

Oy 1 GluArgGlyCysProAlaGluInArgAlaSerProValThrPheGlnAsnGluAspLeu 20
    ::::|||||
Db 1925 GACAAGGCTGCCCGCCGAGCAGAGCCAGCCCTCTGACGTCCCTCGAG----- 1975

Oy 21 GlyProAla---SerProLeuAspSerThr 29
    |||||
Db 1976 GCACCCGCGCTCGCCCGAGCCCGCAGCACA 2005

RESULT 15
US-08-776-251-3
; Sequence 3, Application US/08776251
```

Alignment Scores:			
Pred. No.:	0.0516	Length:	153
Score:	66.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	3
Best Local Similarity:	78.57%	Mismatches:	0
Query Match:	41.25%	Indels:	0
DB:	3	Gaps:	0
JUNC_SEQ8_SEQ4 (1-30) x US-08-776-251-3 (1-153)			
Qy	1	GluArgGlyCysProAlaGluGlnArgAlaSerProValThr	14
Db	10	GACAAGGGGTGCCCGCCGAGCAGAGCAGCGCTCTGACG	51

Search completed: October 16, 2003, 17:16:36  
Job time : 16.6049 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 28.4941 Seconds  
(without alignments)  
2842.104 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ4

Perfect score: 160  
Sequence: 1 ERGCPAEQRAQSPVTFQNEGLGASPDLSTF 30

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q/cn2.1/USPTO\_SPOOL\_P/HOLLERAN480/runat\_15102003\_131912\_20526/app\_query.fasta\_1.4685  
-DB=N.Geneseq\_19Jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=HOLLERAN480 -ACGN 1.1.1758 -runat\_15102003\_131912\_20526 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	147	91.9	2763	24	ABA92252	Mouse Her-2/neu ex
2	147	91.9	2781	24	ABA92253	Her-2/neu extracel
3	83	51.9	1115	21	AAF21778	Human breast and o
4	83	51.9	1755	24	AAD32746	Human CDNA for the
5	83	51.9	1767	24	AAD32744	Human CDNA for the
6	83	51.9	1773	24	AAD32747	Human CDNA for the
7	83	51.9	1806	24	AAD32745	Human CDNA for the
8	83	51.9	3600	21	AAA89736	Human HER-2/neu co
9	83	51.9	3678	24	ABK86207	CDNA encoding huma
10	83	51.9	3768	17	AAT40739	HER-2/neu oncogene
11	83	51.9	3768	20	AAK01912	Human HER-2/neu on
12	83	51.9	3768	21	AAA09455	Human heregulin 2
13	83	51.9	3768	22	AAH23392	Human HER-2/neu pr
14	83	51.9	3768	24	ABZ35744	Human ERBB2 polynu
15	83	51.9	3768	24	ABX09877	Human ERBB2 DNA fr
16	83	51.9	3768	24	AAD43935	Human HER-2 CDNA.
17	83	51.9	3768	24	AAD43986	Human Her2 antigen
18	83	51.9	3768	24	ABV78168	Human ERBB2 DNA SE
19	83	51.9	3768	24	AAD32743	Human Her-2/neu pr
20	83	51.9	3768	24	ABA92250	Human Her-2/neu CD
21	83	51.9	3768	24	ABK10730	Human Her-2/neu DN
22	83	51.9	3768	24	ABL91709	Human polynucleoti
23	83	51.9	3768	24	ABK14058	Human HER2 (Erbb2)
24	83	51.9	4299	14	AAQ46083	Sequence encoding the
25	83	51.9	4472	21	AAA14812	CDNA encoding the
26	83	51.9	4473	19	ABQ76220	Human tumour antiq
27	83	51.9	4473	20	AZ33071	HER-2 nucleic acid
28	83	51.9	4473	24	ABZ34969	Human gene express
29	83	51.9	4473	24	AAD38904	Human Her-2 DNA.
30	83	51.9	4530	16	AAT01585	Her-2/neu (ERBB2/c
31	83	51.9	4530	18	AAT71253	Human Her2 gene.
32	83	51.9	4530	21	AZ60815	Nucleotide sequenc
33	83	51.9	4530	22	AAD19731	Human tyrosine kin
34	83	51.9	4530	24	ABZ35012	Human gene express
35	83	51.9	4530	24	ABV94128	Breast carcinoma r
36	83	51.9	4530	24	ABN85585	Human HER2-neu SQ
37	83	51.9	4530	24	ABK83918	Human CDNA differe
38	83	51.9	4530	25	ACC50139	Breast cancer asso
39	83	51.9	4530	25	ABQ83856	Human Her2/Neu enc
40	83	51.9	9274	22	AAF24297	HER2 transgene pla
41	83	51.9	9274	24	AAD43934	HER-2 transgene pl
42	83	51.9	9274	24	ABK14057	Human HER2 (Erbb2)
43	81	50.6	3771	21	AAA89737	Mouse Her-2/neu CD
44	81	50.6	3771	22	AAA89737	Nucleotide sequenc
45	81	50.6	3771	24	ABA92251	Mouse Her-2/neu CD

ALIGNMENTS

RESULT 1  
ABA92252  
ID ABA92252 standard; CDNA; 2763 BP.  
XX  
AC ABA92252;  
XX  
DT 17-JUN-2002 (first entry)  
XX  
DE Mouse Her-2/neu extracellular-phosphorylation domain fusion cDNA.  
XX  
KW Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;  
XX  
FT receptor; mouse; gene therapy; gene; ss.  
XX  
OS Mus musculus.  
XX  
FH Key  
CDS Location/Qualifiers  
1..2763

```

FT      /*tag= a
XX      /product= "Her-2/neu ECD-PD fusion"
XX
PN      WO200212341-A2.
XX
PD      14-FEB-2002.
XX
PF      03-AUG-2001; 2001WO-US24283.
XX
PR      03-AUG-2000; 2000US-0632507.
XX
PA      (CORI-) CORIXA CORP.
XX      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI      Cheever MA, Gheysen D;
XX
DR      WPI; 2002-241743/29.
XX      P-PSDB; AAM51152.
XX
PT      Her-2/neu fusion protein for treating or preventing cancer by eliciting
PT      or enhancing an immune response to the protein, has Her-2/neu
PT      extracellular domain fused to Her-2/neu intracellular or
PT      phosphorylation domain
XX
PS      Disclosure; Fig 23; 141pp; English.
XX
CC      The present sequence is that of cDNA encoding a fusion between
CC      the extracellular domain (ECD) and the phosphorylation domain (PD)
CC      murine Her-2/neu oncoprotein (see AAM51151). The Her-2/neu gene is
CC      amplified and the oncoprotein is overexpressed in a variety of
CC      human cancers, including breast, ovarian, colon, lung and prostate
CC      cancer. Her-2/neu overexpression correlates with a poor prognosis
CC      in breast and ovarian cancers. The invention provides Her-2/neu
CC      fusion proteins, nucleic acids encoding them, viral vectors, and
CC      vaccines comprising the fusion proteins or nucleic acid molecules.
CC      In preferred fusion proteins, the ECD of Her-2/neu is fused to a
CC      Her-2/neu intracellular domain or PD (or its DeltapD fragment). An
CC      immune response to Her-2/neu protein is elicited or enhanced by
CC      administering the fusion protein in the form of a vaccine, or by
CC      transfecting cells of an animal ex vivo with a nucleic acid
CC      specific T-cells are useful for inhibiting the development of a
CC      cancer, especially breast, ovarian, colon, lung or prostate cancer
CC      in a patient. T cells that specifically react with a Her-2/neu
CC      fusion protein can be used to remove tumour cells from a sample in
CC      order to inhibit the development of cancer in a patient.
XX
SQ      Sequence 2763 BP; 571 A; 855 C; 772 G; 565 T; 0 other;

Alignment Scores:
Pred. No.:      1.96e-11      Length:      2763
Score:          147.00      Matches:      27
Percent Similarity: 96.67%      Conservative: 2
Best Local Similarity: 90.00%      Mismatches: 1
Query Match:      91.88%      Indels:      0
DB:             24      Gaps:      0

JUNC_SEQ8_SEQ4 (1-30) x ABA92252 (1-2763)
Oy      1 GluArgGLyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Db      1918 GAACAGAGCTGCCAGCAGACGAGAGCCAGCCAGCCAGTGTCTCAGACGAGGACTTA 1977

Oy      21 GlyProAlaSerProLeuAspSerThrPhe 30
Db      1978 GGCCCCCTCCAGCCCCATGGACAGCACCTTC 2007

RESULT 2
ABA92253
ID      ABA92253 standard; cDNA; 2781 BP.
XX
AC      ABA92253;

```

```

XX      17-JUN-2002 (first entry)
XX
XX      Her-2/neu extracellular-phosphorylation domain-Tcp0 fusion cDNA.
XX
XX      Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
XX      receptor; Tcp0; mouse; gene therapy; gene; ss.
XX
XX      Chimeric - Mus musculus.
XX      OS      Chimeric - Unidentified.
XX
XX      Key      Location/Qualifiers
XX      CDS      1..2781
XX      FT      /*tag= a
XX      FT      /product= "Her-2/neu ECD-PD-Tcp0 fusion"
XX
XX      WO200212341-A2.
XX
XX      14-FEB-2002.
XX
XX      03-AUG-2001; 2001WO-US24283.
XX
XX      03-AUG-2000; 2000US-0632507.
XX
XX      (CORI-) CORIXA CORP.
XX      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX      Cheever MA, Gheysen D;
XX
XX      WPI; 2002-241743/29.
XX      P-PSDB; AAM51153.
XX
XX      Her-2/neu fusion protein for treating or preventing cancer by eliciting
XX      or enhancing an immune response to the protein, has Her-2/neu
XX      extracellular domain fused to Her-2/neu intracellular or
XX      phosphorylation domain
XX
XX      Disclosure; Fig 25; 141pp; English.
XX
CC      The present sequence is that of cDNA encoding a fusion between
CC      the extracellular domain (ECD) and the phosphorylation domain (PD)
CC      murine Her-2/neu oncoprotein (see AAM51151) plus a C-terminal
CC      Tcp0 motif that improves immunogenicity. The Her-2/neu gene is
CC      amplified and the oncoprotein is overexpressed in a variety of
CC      human cancers, including breast, ovarian, colon, lung and prostate
CC      cancer. Her-2/neu overexpression correlates with a poor prognosis
CC      in breast and ovarian cancers. The invention provides Her-2/neu
CC      fusion proteins, nucleic acids encoding them, viral vectors, and
CC      vaccines comprising the fusion proteins or nucleic acid molecules.
CC      In preferred fusion proteins, the ECD of Her-2/neu is fused to a
CC      Her-2/neu intracellular domain or PD (or its DeltapD fragment). An
CC      immune response to Her-2/neu protein is elicited or enhanced by
CC      administering the fusion protein in the form of a vaccine, or by
CC      transfecting cells of an animal ex vivo with a nucleic acid
CC      specific T-cells are useful for inhibiting the development of a
CC      cancer, especially breast, ovarian, colon, lung or prostate cancer
CC      in a patient. T cells that specifically react with a Her-2/neu
CC      fusion protein can be used to remove tumour cells from a sample in
CC      order to inhibit the development of cancer in a patient.
XX
SQ      Sequence 2781 BP; 574 A; 859 C; 779 G; 569 T; 0 other;

Alignment Scores:
Pred. No.:      1.97e-11      Length:      2781
Score:          147.00      Matches:      27
Percent Similarity: 96.67%      Conservative: 2
Best Local Similarity: 90.00%      Mismatches: 1
Query Match:      91.88%      Indels:      0
DB:             24      Gaps:      0

JUNC_SEQ8_SEQ4 (1-30) x ABA92253 (1-2781)

```

Qy	1	GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu	20
Db	1918	GAACGAGGCTGCCAGCAGACGAGAGCCAGCCAGCTCTCAGAACGAGACTTA	1977
Qy	21	GlyProAlaSerProLeuAspSerThrPhe	30
Db	1978	GGCCCTCCAGCCCATGGACGACCCTTC	2007
RESULT 3			
AAF21778			
ID	AAF21778	standard; DNA: 1115 BP.	
XX	AAF21778;		
XX	27-MAR-2001	(first entry)	
XX	Human breast and ovarian cancer associated antigen gene	SEQ ID 165.	
XX	Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;		
KW	neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;		
KW	antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;		
KW	antibacterial; antifungal; antiparasitic; cardiant; immune disorder;		
KW	Addison's disease; allergy; autoimmune haemolytic anaemia;		
KW	autoimmune thyroiditis; diabetes mellitus; Crohn's disease;		
KW	multiple sclerosis; rheumatoid arthritis; ulcerative colitis;		
KW	cardiovascular disorder; wound healing; neurological disease; ds.		
XX	Homo sapiens.		
OS	WO200055173-A1.		
PN	21-SEP-2000.		
XX	08-MAR-2000; 2000WO-US05881.		
XX	12-MAR-1999; 99US-0124270.		
XX	(HUMA-) HUMAN GENOME SCI INC.		
XX	Rosen CA, Ruben SM;		
PI	WPI; 2000-611515/58.		
XX	P-PSDB; AAB56875.		
DR	New human breast and ovarian cancer associated gene sequences and the		
PT	polypeptides encoded by these genes, useful in the prevention,		
PT	treatment and diagnosis of cancer, immune disorders, cardiovascular		
PT	disorders and neurological diseases -		
XX	Claim 1; Page 604; 1299pp; English.		
PS	Sequences AAF21614 - AAF22031 represent DNA sequences encoding human		
XX	proteins AAB58711 - AAB59128. The DNA and protein sequences are		
CC	associated with breast and ovarian cancer. Included in the invention are		
CC	sequences AAF22032 - AAF22040 and AAB59129 which are used in the		
CC	isolation and characterisation of the DNA and protein sequences of the		
CC	invention. The breast and ovarian cancer associated DNA, protein, agonist		
CC	or antagonist sequences exhibit cytostatic; immunosuppressive;		
CC	neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;		
CC	antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;		
CC	antibacterial; antifungal; antiparasitic and cardiant activity. The		
CC	polynucleotide and protein sequences are used in the diagnosis of cancer,		
CC	particularly breast and ovarian cancer. The nucleic acid sequences,		
CC	proteins, agonists and antagonists may also be used in the diagnosis,		
CC	prevention and treatment of immune disorders e.g. Addison's disease,		
CC	allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,		
CC	diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid		
CC	arthritis and ulcerative colitis; cardiovascular disorders such as		
CC	myocardial ischaemias; wound healing; neurological diseases such as		
CC	cerebral anoxia and epilepsy; and infectious diseases.		
XX	Sequence 1115 BP; 210 A; 336 C; 338 G; 222 T; 9 other;		
XX			

The composition is useful for the therapy and diagnosis of cancer, preferably breast cancer, in pharmaceutical compositions, e.g., vaccine and other compositions for the diagnosis, prevention and treatment of human malignancies, for stimulating and/or expanding T cells specific for Her-2/Neu polypeptide and for inhibiting the development of cancer in a patient. The invention is useful for stimulating a T cell response in a human patient, as probe or primer for nucleic acid hybridisation, to selectively form duplex molecules with complementary stretches of the entire Her-2/Neu gene or gene fragments of interest, to isolate a full length gene from a suitable library, and to direct expression of a polypeptide in appropriate host cells. The composition is useful in prophylactic or therapeutic applications and for the treatment of cancer, preferably for the immunotherapy of breast cancer and other Her-2/Neu-associated malignancies. The invention is useful in gene therapy. The present sequence is human cDNA for the clone HTCD CT.His. coding region.

CC . polypeptide in appropriate host cells. The composition is useful in  
CC prophylactic or therapeutic applications and for the treatment of cancer,  
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-  
CC associated malignancies. The invention is useful in gene therapy. The  
CC present sequence is human cDNA for the clone HICD\_CT\_His\_coding\_region.  
xx  
SQ sequence 1767 BP; 381 A; 521 C; 529 G; 336 T; 0 other;

Score:	83.00	Matches:	16
Percent similarity:	88.80%	Conservative:	0

Best Local Similarity:	88.89%	Mismatches:	2
Query Match:	51.98%	Indels:	0

DB:

THING CUB8 C608 /1-20\ " X8D3C7A /1-1757\

JUNC\_\_SEQ8\_\_SEQ4 (1-30) X AAD3Z/44 (1-17)

13  
QY

D<sub>b</sub> 937 GTGGTCATCCAGAAATGA

0  
1  
2  
3  
4  
5  
6  
7  
8  
9  
A  
B  
C  
D  
E  
F  
G  
H  
I  
J  
K  
L  
M  
N  
O  
P  
Q  
R  
S  
T  
U  
V  
W  
X  
Y  
Z

## RESULT 6

AAD32747

10 AAD32747 Standard, CERN, 1773 BE.  
XX  
AC AAD32747;

XX  
DT  
01-JUL-2002

Human cDNA for the clone H

XX  
XX  
DE  
HUMAN CDNA FOR THE C

KW Human: Her-2/Neu pro

KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.  
XX

OS Homo sapiens.

EH	Key	Location/Qualif
EH	CDS	1 1770

FT	/tag= a	Human protein encoded by cDNA for the clone
FT	/product=	

FT  
XX  
HICD\_in\_ppdm

XX  
PN  
WQ200214503-A2

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

PD 21-FEB-2002.

XX

PF 14-AUG-2001; 2001WO-US41733.

XX

PR 14-AUG-2000; 2000US-225152P.

28-SEP-2000; 2000US-236428P.

PR 21-FEB-2001; 200105-270520P.

XX  
XX  
DN  
(CONT - )  
CORPUS CORP

PA (CORI-) CORIAX CORP.

Hand-zimmermann & Cheever  
XX

PT McNeill PD Vedvyick TS.  
PT Manda Zimmerman S, Cheever

'CT NOTADNA /D/ TTTCOTY  
XX

DR WPT: 2002-280758/32

DR P-PSDB: AAE20484.

XX

PT Novel isolated Her-2/Neu pol

PT prevention and diagnosis of

XX PS Example 5; Page 119; 129pp; English.

XX PI The invention relates to an isolated Her-2/Neu polypeptide composition

XX CC effective for eliciting an immune response. The invention is useful for

CC eliciting an immune response in a patient, where the patient is human

CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.

CC The composition is useful for the therapy and diagnosis of cancer,

CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine

CC and other compositions for the diagnosis, prevention and treatment of

CC human malignancies, for stimulating and/or expanding T cells specific for

CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a

CC patient. The invention is useful for stimulating a T cell response in a

CC human patient, as probe or primer for nucleic acid hybridisation, to

CC selectively form duplex molecules with complementary stretches of the

CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full

CC length gene from a suitable library, and to direct expression of a

CC polypeptide in appropriate host cells. The composition is useful in

CC prophylactic or therapeutic applications and for the treatment of cancer,

CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-

CC associated malignancies. The invention is useful in gene therapy. The

CC present sequence is human cDNA for the clone HICD\_in\_pPDM\_coding\_

XX CC sequence.

XX SQ Sequence 1773 BP; 383 A; 528 C; 530 G; 332 T; 0 other;

Alignment Scores:

Pred. No.:	0.0242	Length:	1773
Score:	83.00	Matches:	16
Percent Similarity:	88.8%	Conservative:	0
Best Local Similarity:	88.8%	Mismatches:	2
Query Match:	51.88%	Indels:	0
DB:	24	Gaps:	0

JUNC\_SEQ8\_SEQ4 (1-30) x AAD32747 (1-1773)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 961 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 1014

RESULT 7

AAD32745

ID AAD32745 standard; cDNA; 1806 BP.

XX AC AAD32745;

XX DT 01-JUL-2002 (first entry)

XX DE Human cDNA for the clone HICD\_plus\_8\_HIS.

XX KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;

XX KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 1..1803

FT /\*tag= a

FT /product= "Human protein encoded by cDNA for the clone

FT HICD\_plus\_8\_HIS"

FT /transl\_except= (pos:1543..1545, aa:Pro)

XX PN WO200214503-A2.

XX PD 21-FEB-2002.

XX PF 14-AUG-2001; 2001WO-US41733.

XX PR 14-AUG-2000; 2000US-225152P.

PR 28-SEP-2000; 2000US-236428P.

PR 21-FEB-2001; 2001US-270520P.

XX PA (CORI-) CORIXA CORP.

XX PI Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;

XX PI McNeill PD, Vedrick TS;

XX DR WPI; 2002-280758/32.

XX DR P-PSDB; AAE20482.

XX PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,

XX CC prevention and diagnosis of cancer, preferably breast cancer

PS Example 5; Page 118; 129pp; English.

XX CC The invention relates to an isolated Her-2/Neu polypeptide composition

CC effective for eliciting an immune response. The invention is useful for

CC eliciting an immune response in a patient, where the patient is human

CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.

CC The composition is useful for the therapy and diagnosis of cancer,

CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine

CC and other compositions for the diagnosis, prevention and treatment of

CC human malignancies, for stimulating and/or expanding T cells specific for

CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a

CC patient. The invention is useful for stimulating a T cell response in a

CC human patient, as probe or primer for nucleic acid hybridisation, to

CC selectively form duplex molecules with complementary stretches of the

CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full

CC length gene from a suitable library, and to direct expression of a

CC polypeptide in appropriate host cells. The composition is useful in

CC prophylactic or therapeutic applications and for the treatment of cancer,

CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-

CC associated malignancies. The invention is useful in gene therapy. The

CC present sequence is human cDNA for the clone HICD\_plus\_8\_HIS.

XX SQ Sequence 1806 BP; 391 A; 530 C; 544 G; 341 T; 0 other;

Alignment Scores:

Pred. No.:	0.0248	Length:	1806
Score:	83.00	Matches:	16
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	2
Query Match:	51.88%	Indels:	0
DB:	24	Gaps:	0

JUNC\_SEQ8\_SEQ4 (1-30) x AAD32745 (1-1806)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 994 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 1047

RESULT 8

AAA89736

ID AAA89736 standard; DNA; 3600 BP.

XX AC AAA89736;

XX DT 12-JAN-2001 (first entry)

XX DE Human HER-2/neu coding sequence.

XX KW Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;

XX KW breast cancer; prostate cancer; ovarian cancer; lung cancer;

XX KW colon cancer; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 1..3600

FT /\*tag= a

FT /product= "HER-2/neu protein"

XX PN WO2000044899-A1.

XX PD 03-AUG-2000.

```

PF 28-JAN-2000; 2000WO-US02164.
XX
PR 29-JAN-1999; 99US-0117976.
XX
PA (CORI-) CORIXA CORP.
PA (SMK ) SMITHKLINE BEECHAM.
XX
PI Cheever MA, Gheysen D;
XX
DR WPI; 2000-505976/45.
DR P-PSDB; RAB21198, RAB21208.
XX
PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and
PT prostate cancers -
XX
PS Disclosure; Fig 15; 128pp; English.
XX
CC The present sequence encodes the human HER-2/neu protein. HER-2/neu is
CC a member of the tyrosine kinase family of receptor-like glycoproteins
CC and shows homology to the epidermal growth factor receptor (EGFR). It
CC probably plays a part in cell growth and/or differentiation. The
CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising
CC a HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation
CC domain may be used to treat or prevent cancer by eliciting or
CC enhancing an immune response to the HER-2/neu protein. It may be used
CC to treat malignancies such as breast, ovarian, colon, lung and
CC prostate cancers, and may be used as an antigen to vaccinate against
CC these neoplasias.
XX
SQ Sequence 3600 BP; 723 A; 1108 C; 1075 G; 694 T; 0 other;

Alignment Scores:
Pred. No.: 0.0566 Length: 3600
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 21 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AAB89736 (1-3600)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGACCTTC 3012

RESULT 9
ABK86207
ID ABR86207 standard; cDNA; 3678 BP.
XX
AC ABR86207;
XX
DT 24-SEP-2002 (first entry)
XX
DE cDNA encoding human breast cancer antigen, Her2 variant.
XX
KW Human; Her2; cytostatic; antiviral; immunostimulant;
KW cell-mediated immune response; tumour; breast cancer;
KW virus infection; prostate cancer; colorectal cancer; pancreatic cancer;
KW lymphoma; leukaemia; hepatitis; lentivirus; herpesvirus;
KW human immunodeficiency virus; HIV; flavivirus; pestivirus; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 7..3678
FT a
FT /product= "Breast cancer antigen Her2 variant"
XX
PN W0200240059-A2.
XX
PD 23-MAY-2002.
XX

01-NOV-2001; 2001WO-US45626.
XX
PR 01-NOV-2000; 2000US-0704232.
XX
PA (AMBI-) AMERICAN FOUND BIOLOGICAL RES INC.
PA (MINC/) MINCHEFF M S.
PA (LOUK/) LOUKINOV D I.
PA (ZOUB/) ZOUBAK S.
XX
PI Mincheff MS, Loukinov DI, Zoubak S;
XX
DR WPI; 2002-527524/56.
DR P-PSDB; AAU98923.
XX
PT Inducing a cell-mediated immune response against a target antigen,
PT reducing undesired cells and stimulating presentation of an antigen by
PT a cell, comprises administering a polynucleotide encoding a variant of
PT an antigen -
XX
PS Disclosure; Page 128-134; 146pp; English.
XX
CC The invention relates to a method of inducing a cell-mediated immune
CC response against a cell comprising a target antigen (I) in a subject,
CC treating a subject having undesired cells, for example tumour cells
CC or virally infected cells (C), reducing the number of (C) in a subject,
CC and stimulating presentation of (I) by a cell. This is done by
CC administering a polynucleotide (II) encoding a variant of (I), so that
CC (II) expressed in a cell and cell-mediated immune response is induced.
CC The method can be used to treat prostate cancer, breast cancer,
CC colorectal cancer and pancreatic cancer, as well as lymphomas and
CC leukaemias. The method is also useful in treating chronic viral
CC infections such as those caused by hepatitis viruses, lentiviruses
CC (including human immunodeficiency virus (HIV)), herpesviruses and the
CC flaviviruses and pestiviruses. The present sequence represents the coding
CC sequence of human breast cancer antigen, Her2 variant, used as a target
CC antigen in the method of the invention.
XX
SQ Sequence 3678 BP; 746 A; 1137 C; 1091 G; 704 T; 0 other;

Alignment Scores:
Pred. No.: 0.0581 Length: 3678
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 24 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x ABR86207 (1-3678)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2869 GTGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGACCTTC 2922

RESULT 10
AAT40739
ID AAT40739 standard; cDNA; 3768 BP.
XX
AC AAT40739;
XX
DT 01-JAN-1997 (first entry)
XX
DE HER-2/neu oncogene.
XX
KW HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;
KW breast cancer; ovary cancer; colon cancer; lung cancer;
KW prostate cancer; genetic immunisation; tumour; vaccine; vector;
KW ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3765
FT b
FT /*tag= b

```

```

FT      /note= "nucleotides 2026-3765 (claim 1) code for
FT      HER-2/neu intracellular domain"
XX      WO9630514-A1.
PN      03-OCT-1996.
XX      28-MAR-1996; 96WO-US01689.
XX      31-MAR-1995; 95US-0414417.
XX      (UNIW ) UNIV WASHINGTON.
XX      Cheever MA, Disis ML;
XX      WPI; 1996-455361/45.
XX      P-PSDB; AAW01111.
XX      DNA encoding HER-2-neu poly:peptide(s) - used for prevention or
XX      treatment of malignancies with which the HER-2/neu oncogene is
XX      associated
XX      Claim 1; Page 49-56; 71pp; English.
XX      Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or
XX      c-erbB2) protein (AAW01111). The oncogene is overexpressed in various
XX      cancers, including breast, ovarian, colon, lung and prostate, and
XX      appears to induce malignancies through quantitative mechanisms that
XX      result from increased or deregulated expression of an essentially
XX      normal gene product. Nucleotides 2026-3765 of the cDNA sequence
XX      code for the intracellular domain (Lys676-Val1255) of the HER-2/neu
XX      protein, which is useful for immunisation against malignancy.
XX      Nucleic acids can be used to direct expression of the intracellular
XX      domain in transformed host cells, or are used, alone or in a viral
XX      vector, for genetic immunisation of an animal.
XX      SQ      Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.:      0.0598      Length:      3768
Score:          83.00      Matches:      16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:      51.88%      Indels:      0
DB:              17      Gaps:      0

JUNC_SEQ8_SEQ4 (1-30) x AAT40739 (1-3768)
QY      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
      ||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3012

RESULT 11
AAAX01912
ID      AAX01912 standard; DNA; 3768 BP.
XX
AC      AAX01912;
XX
DT      21-APR-1999 (first entry)
XX
DE      Human HER-2/neu oncogene DNA.
XX
KW      HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
KW      malignancy; treatment; tumour; ss.
XX
OS      Homo sapiens.
XX
FH      Key      Location/Qualifiers
CDS      1..3768
FT      /*tag= a
FT      /product= "HER-2/neu"
FT      /note= "oncogene"
FT      misc_feature      2026..3765

```

```

FT      /*tag= b
FT      /note= "region which elicits immune response"
XX
PN      US5869445-A.
XX
PD      09-FEB-1999.
XX
PF      01-APR-1996; 96US-0625101.
XX
PR      01-APR-1996; 96US-0625101.
PR      17-MAR-1993; 93US-0033644.
PR      12-AUG-1993; 93US-0106112.
PR      31-MAR-1995; 95US-0414417.
XX
PA      (UNIW ) UNIV WASHINGTON.
XX
PI      Cheever MA, Disis ML;
XX
PI      WPI; 1999-152835/13.
XX      P-PSDB; AAW92406.
XX
PT      Use of HER-2/neu polypeptides - for eliciting an immune response to
PT      an HER-2/neu associated malignancy, particularly for treating or
PT      preventing tumours
XX
PS      Claim 1a; Column 23-32; 26pp; English.
XX
CC      This sequence encodes the human HER-2/neu oncogene protein. A fragment
XX      of this protein is used in a method for eliciting or enhancing an immune
XX      response to HER-2/neu protein. The polypeptide can stimulate T cells and
XX      B cells to produce an immune response to the HER-2/neu protein. The
XX      method can be used for immunisation against a malignancy in which the
XX      HER-2/neu oncogene is associated and in the treatment of an existing
XX      tumour, or to prevent tumour occurrence or reoccurrence.
XX
SQ      Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.:      0.0598      Length:      3768
Score:          83.00      Matches:      16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:      51.88%      Indels:      0
DB:              20      Gaps:      0

JUNC_SEQ8_SEQ4 (1-30) x AAX01912 (1-3768)
QY      13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
      ||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGTCCCTTGGACAGCACCTTC 3012

RESULT 12
AAAX09455
ID      AAX09455 standard; DNA; 3768 BP.
XX
AC      AAX09455;
XX
DT      10-AUG-2000 (first entry)
XX
DE      Human heregulin 2 (Her2) coding sequence.
XX
KW      Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
KW      self-protein; cell-associated peptide antigen; foreign epitope;
KW      cancer; breast cancer; prostate cancer; ss.
XX
OS      Homo sapiens.
XX
PN      WO200020027-A2.
XX
PD      13-APR-2000.
XX
PF      05-OCT-1999; 99WO-DK00525.
XX

```

```

PR 05-OCT-1998; 98DK-0001261.
PI 20-OCT-1998; 98US-0105011.
XX
XX (MEBI-) M & E BIOTECH AS.
XX
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
XX WPI; 2000-349917/30.
DR P-PSDB; AAY92620.
XX
XX Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
XX Claim 62; Page 187-193; 220pp; English.
XX
XX The claims detail a method for inducing immune responses against weakly
CC immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (i.e. self-proteins), for example, human
CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively.
XX
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 21 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AAA09455 (1-3768)
Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGACCTTC 3012

RESULT 13
AAH23392
ID AAH23392 standard; DNA; 3768 BP.
XX
XX AAH23392;
XX
XX 25-SEP-2001 (first entry)
XX
XX Human HER-2/neu protein encoding DNA.
XX
XX Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
KW oncogene; cancer; cytostatic; vaccine; pl85; c-erbB2; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 1..3768
FT CDS
FT /*tag= a
FT /product= "HER-2/neu protein"
XX
XX WO200153463-A2.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US01850.

```

```

XX 21-JAN-2000; 2000US-0177545.
XX
XX (CORI-) CORIXA CORP.
XX
XX Cheever MA, Hand-Zimmermann S;
XX
XX WPI; 2001-476112/51.
DR P-PSDB; AAB85458.
XX
XX New antigen-presenting cells, useful as vaccines for eliciting or
PT enhancing an immune response to HER-2/neu protein, particularly useful
PT for treating or preventing cancer, e.g. breast cancer
XX
XX Claim 1; Page 41-46; 49pp; English.
XX
XX The invention provides an isolated antigen-presenting cell, which
CC expresses at least an immunogenic portion of a polypeptide that produces
CC an immune response to HER-2/neu protein. The antigen-presenting cells are
CC useful as vaccines for eliciting or enhancing an immune response to
CC HER-2/neu protein, particularly in treating or preventing malignancies in
CC which the HER-2/neu oncogene is associated. Specifically, these are
CC useful for treating or preventing cancer, e.g. breast cancer, ovarian,
CC colon, lung or prostate cancers. The present sequence represents a DNA
CC encoding the human HER-2/neu protein (also known as p185 or c-erbB2).
XX
XX Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 0.0598 Length: 3768
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 22 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AAH23392 (1-3768)
Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 2959 GTGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGACCTTC 3012

RESULT 14
ABZ35744
ID ABZ35744 standard; DNA; 3768 BP.
XX
XX AC ABZ35744;
XX
XX 07-FEB-2003 (first entry)
XX
XX Human ERBB2 polynucleotide SEQ ID NO 52.
XX
XX Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
KW protozoacide; gene expression; antisense; tumour; infection; Plasmodium;
KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
KW Hepatitis C virus; human papilloma virus; gene; ds.
XX
XX Homo sapiens.
XX
XX PN DE10100588-A1.
XX
XX 18-JUL-2002.
XX
XX 09-JAN-2001; 2001DE-1000588.
XX
XX 09-JAN-2001; 2001DE-1000588.
XX
XX (RIBO-) RIBOPHARMA AG.
XX
XX Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
XX WPI; 2002-683450/74.

```



PT Inhibiting expression of target genes, useful e.g. for treating tumors,  
PT by introducing into cells two double-stranded RNAs that are  
PT complementary to the target -  
XX  
XX Claim 13; Page 38-39; 100pp; German.  
XX  
XX The invention relates to inhibiting expression of a target gene in a cell  
CC by introducing at least two oligoribonucleotides (dsRNAI and II), both  
CC with a double-stranded (ds) structure of at most 49 sequential nucleotide  
CC pairs. At least part of one strand (S1, S2) of the ds structures in each  
CC of dsRNAI and II are complementary to regions in the target gene. The  
CC method uses antisense inhibition of gene expression using double stranded  
CC RNA inhibition (RNAi). The method is particularly used to treat tumours  
CC or infections, especially by Plasmodium or viruses/viroids (pathogenic on  
CC humans, animals or plants). The method provides more effective inhibition  
CC of expression than known methods using a single dsRNA, even at very low  
CC concentrations. When dsRNA has at least one unpaired nucleotide at the  
CC end, stability (and thus effective concentration in the cell) is  
CC improved and efficiency can be increased further by pretreating the cells  
CC with interferon. The present sequence is that of a target DNA of the  
CC invention.  
XX  
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 0.0598 Length: 3768  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 24 Gaps: 0  
  
JUNC\_SEQ8\_SEQ4 (1-30) x ABZ35744 (1-3768)  
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 2959 GTGGTCATCCAGAAATGAGGACTTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 3012  
  
RESULT 15  
ABX09987  
ID ABX09987 standard; DNA; 3768 BP.  
XX  
XX AC ABX09987;  
XX  
XX DT 23-JAN-2003 (first entry)  
XX  
XX DE Human ERBB2 DNA fragment SEQ ID 52.  
XX  
XX KW Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;  
XX KW prion; inhibition; human; ds.  
XX  
XX OS Homo sapiens.  
XX  
XX PN DE10100587-C1.  
XX  
XX PD 21-NOV-2002.  
XX  
XX PF 09-JAN-2001; 2001DE-1000587.  
XX  
XX PR 09-JAN-2001; 2001DE-1000587.  
XX  
XX PA (RIBO-) RIBOPHARMA AG.  
XX  
XX PI Kreutzer R, Limmer S, Rost S, Hadwiger P;  
XX  
XX DR WPI; 2002-742209/81.  
XX  
XX PT Inhibiting expression of target genes, e.g. oncogenes, in cells, by  
XX PT introduction of complementary double-stranded oligoribonucleotide,  
XX PT after treating the cell with interferon -  
XX  
XX PS Disclosure: Page 43-44; 98pp; German.  
XX

CC This invention describes a novel method for inhibiting expression of a  
CC target gene by introducing into the cell that contains the target gene  
CC at least one oligoribonucleotide (dsRNAI) that has a double-stranded  
CC (ds) structure of not more than 49 consecutive nucleotides (nt), where  
CC at least a segment of one strand of the ds structure is complementary  
CC with the target gene and the cells are treated with interferon before  
CC introduction of dsRNAI. The method is used to inhibit expression of  
CC target genes, particularly oncogenes, cytokine genes, Id (not defined)  
CC protein genes; developmental or prion genes, or genes expressed in  
CC pathogenic organisms (particularly plasmodia) or in viruses or viroids  
CC (pathogenic in humans, animals or plants). Treating the cells with  
CC interferon greatly increases the extent to which dsRNA can inhibit  
CC expression of the target genes, and the effect is even greater when dsRNA  
CC are modified to increase their stability. ABX09936-ABX10075 represent  
CC gene fragments used to illustrate the method of the invention.  
XX  
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 0.0598 Length: 3768  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 24 Gaps: 0  
  
JUNC\_SEQ8\_SEQ4 (1-30) x ABX09987 (1-3768)  
QY 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 2959 GTGGTCATCCAGAAATGAGGACTTTGGGCCAGCCAGCTCCCTTGGACAGCACCTTC 3012  
  
Search completed: October 15, 2003, 23:48:48  
Job time : 31.4941 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:09:13 : Search time 432.026 Seconds  
(without alignments)  
2840.777 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ4  
Perfect score: 160  
Sequence: 1 ERGCPAEQRASPVTFQEDLGASPLDSTF 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/HOLLERAM480/runat\_15102003\_131912\_20536/app\_query.fasta\_1.4685  
-DB=GenEmbl -OFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAM480 -CGN\_1\_1\_22724.@runat\_15102003\_131912\_20536 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg\_hum.\*  
31: em.htg\_inv.\*  
32: em.htg\_other.\*  
33: em.htg\_mus.\*  
34: em.htg\_pln.\*  
35: em.htg\_rod.\*  
36: em.htg\_mam.\*  
37: em.htg\_vrt.\*  
38: em.sy.\*  
39: em.htgo\_hum.\*  
40: em.htgo\_mus.\*  
41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	147	91.9	2763	6	AX380942	AX380942 Sequence
2	147	91.9	2781	6	AX380944	AX380944 Sequence
3	83	51.9	1755	6	AX384609	AX384609 Sequence
4	83	51.9	1767	6	AX384607	AX384607 Sequence
5	83	51.9	1773	6	AX384610	AX384610 Sequence
6	83	51.9	1806	6	AX384608	AX384608 Sequence
7	83	51.9	3678	6	AX505114	AX505114 Sequence
8	83	51.9	3768	6	AR034479	AR034479 Sequence
9	83	51.9	3768	6	AX060704	AX060704 Sequence
10	83	51.9	3768	6	AX201817	AX201817 Sequence
11	83	51.9	3768	6	AX380923	AX380923 Sequence
12	83	51.9	3768	6	AX384604	AX384604 Sequence
13	83	51.9	3768	6	AX465456	AX465456 Sequence
14	83	51.9	3768	6	AX467229	AX467229 Sequence
15	83	51.9	3768	6	AX481438	AX481438 Sequence
16	83	51.9	3780	4	AB008451	AB008451 Canis fam
17	83	51.9	4473	6	AR080259	AR080259 Sequence
18	83	51.9	4473	6	AR167390	AR167390 Sequence
19	83	51.9	4473	9	HSEB2R	X03363 Human c-erb
20	83	51.9	4530	6	AR202597	AR202597 Sequence
21	83	51.9	4530	6	AR283481	AR283481 Sequence
22	83	51.9	4530	6	AX282577	AX282577 Sequence
23	83	51.9	4530	6	AX587649	AX587649 Sequence
24	83	51.9	4530	6	AX644071	AX644071 Sequence
25	83	51.9	4530	6	BD005474	BD005474 Cellular
26	83	51.9	4530	6	I21124	I21124 Sequence 9
27	83	51.9	4530	6	I59745	I59745 Sequence 9
28	83	51.9	4530	9	HUMHER2A	M11730 Human tyros
29	83	51.9	9274	6	AX060703	AX060703 Sequence
30	81	50.6	3771	6	AX189662	AX189662 Sequence
31	81	50.6	3771	6	AX380925	AX380925 Sequence
32	81	50.6	3955	6	AX380924	AX380924 Sequence
33	81	50.6	3955	6	I21129	I21129 Sequence 14
34	81	50.6	3955	6	I59750	I59750 Sequence 14
35	81	50.6	3955	10	RNNEUR	X03362 Rat mRNA fo
36	81	50.6	4694	10	BC046811	BC046811 Mus muscu
37	81	50.6	4695	10	BC053078	BC053078 Mus muscu
38	81	50.6	4727	10	AY116182	AY116182 Rattus no
39	80	50.0	4062	10	HANNEU	D16295 Mesocricetu
40	79	49.4	13450	9	AB096612	AB096612 Homo sapi
41	79	49.4	16572	9	AB096614	AB096614 Homo sapi
42	79	49.4	20271	9	AB096613	AB096613 Homo sapi
43	79	49.4	30837	9	AY208911	AY208911 Homo sapi
44	79	49.4	161815	9	AC079199	AC079199 Homo sapi
45	79	49.4	168585	9	AC040933	AC040933 Homo sapi

ALIGNMENTS

RESULT 1

```
AX380942
LOCUS AX380942 2763 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 28 from Patent WO212341.
ACCESSION AX380942
VERSION AX380942.1 GI:19575786
KEYWORDS
SOURCE synthetic construct
ORGANISM Her-2/neu fusion proteins
REFERENCE 1 Cheever, M.A. and Gheysen, D.
AUTHORS Her-2/neu fusion proteins
TITLE Patent: WO 0212341-A 28 14-FEB-2002;
JOURNAL CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
source
1..2763
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mouse ECD-PD fusion protein cDNA"
BASE COUNT 571 a 855 c 772 g 565 t
ORIGIN
Alignment Scores:
Pred. No.: 3.46e-11 Length: 2763
Score: 147.00 Matches: 27
Percent Similarity: 96.67% Conservative: 2
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 91.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AX380942 (1-2763)
Qy 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCAGCAGACAGAGCCAGCCAGTGCAGTCTCAGACGAGGACTTA 1977

RESULT 2
AX380944
LOCUS AX380944 2781 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 30 from Patent WO212341.
ACCESSION AX380944
VERSION AX380944.1 GI:19575787
KEYWORDS
SOURCE synthetic construct
ORGANISM Her-2/neu fusion proteins
REFERENCE 1 Cheever, M.A. and Gheysen, D.
AUTHORS Her-2/neu fusion proteins
TITLE Patent: WO 0212341-A 30 14-FEB-2002;
JOURNAL CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
source
1..2781
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mouse ECD-PD-TcP0 fusion protein cDNA"
BASE COUNT 574 a 859 c 779 g 569 t
ORIGIN
Alignment Scores:
Pred. No.: 3.49e-11 Length: 2781
Score: 147.00 Matches: 27
Percent Similarity: 96.67% Conservative: 2
Best Local Similarity: 90.00% Mismatches: 1
Query Match: 91.88% Indels: 0
DB: 6 Gaps: 0
```

```
JUNC_SEQ8_SEQ4 (1-30) x AX380944 (1-2781)
Qy 1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCAGCAGACAGAGCCAGCCAGTGCAGTCTCAGACGAGGACTTA 1977

RESULT 3
AX384609
LOCUS AX384609 1755 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 6 from Patent WO0214503.
ACCESSION AX384609
VERSION AX384609.1 GI:19577810
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
McNeill, P.D. and Vedvick, T.S.
JOURNAL Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
Patent: WO 0214503-A 6 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
1..1755
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 376 a 517 c 531 g 331 t
ORIGIN
Alignment Scores:
Pred. No.: 0.0303 Length: 1755
Score: 83.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 51.88% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ8_SEQ4 (1-30) x AX384609 (1-1755)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30
Db 937 GTGGTCATCCAGAAATGAGGACTTGGCCCGCCAGTCCCTTGGACAGACCTTC 990

RESULT 4
AX384607
LOCUS AX384607 1767 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 4 from Patent WO0214503.
ACCESSION AX384607
VERSION AX384607.1 GI:19577808
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,
McNeill, P.D. and Vedvick, T.S.
JOURNAL Compositions and methods for the therapy and diagnosis of
her-2/neu-associated malignancies
Patent: WO 0214503-A 4 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
1..1767
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

BASE COUNT 381 a 521 c 529 g 336 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.0305 Length: 1767  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 6 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x AX384607 (1-1767)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 937 GTGGTCATCCAGAAAGAGGACTTGGGCCAGCCAGCCTTCCTTGGACAGCACCTTC 990

RESULT 5  
AX384610 1773 bp DNA linear PAT 19-MAR-2002  
LOCUS  
DEFINITION Sequence 7 from Patent WO214503.  
ACCESSION AX384610  
VERSION AX384610.1 GI:19577811

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,

McNeill, P.D. and Vedvick, T.S.

Compositions and methods for the therapy and diagnosis of

her-2/neu-associated malignancies

Patient: WO 0214503-A 7 21-FEB-2002;

JOURNAL CORIXA CORPORATION (US)

FEATURES

Location/Qualifiers

1..1773

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT 383 a 528 c 530 g 332 t

ORIGIN

Alignment Scores:  
Pred. No.: 0.0306 Length: 1773  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 6 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x AX384610 (1-1773)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 961 GTGGTCATCCAGAAAGAGGACTTGGGCCAGCCAGCCTTCCTTGGACAGCACCTTC 1014

RESULT 6  
AX384608 1806 bp DNA linear PAT 19-MAR-2002  
LOCUS  
DEFINITION Sequence 5 from Patent WO214503.  
ACCESSION AX384608  
VERSION AX384608.1 GI:19577809

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE Hand-Zimmermann, S., Cheever, M.A., Foy, T.M., Lodes, M.J., Kalos, M.D.,

McNeill, P.D. and Vedvick, T.S.

Compositions and methods for the therapy and diagnosis of

her-2/neu-associated malignancies

JOURNAL Patent: WO 0214503-A 5 21-FEB-2002;  
CORIXA CORPORATION (US)

FEATURES

Location/Qualifiers

1..1806

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT 391 a 530 c 544 g 341 t

ORIGIN

Alignment Scores:

Pred. No.: 0.0312 Length: 1806

Score: 83.00 Matches: 16

Percent Similarity: 88.89% Conservative: 0

Best Local Similarity: 88.89% Mismatches: 2

Query Match: 51.88% Indels: 0

DB: 6 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x AX384608 (1-1806)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30

Db 994 GTGGTCATCCAGAAAGAGGACTTGGGCCAGCCAGTCTTGGACAGCACCTTC 1047

RESULT 7

AX505114

LOCUS

DEFINITION Sequence 7 from Patent WO240059.

ACCESSION AX505114

VERSION AX505114.1 GI:23386421

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS Mincheff, M.S., Loukinov, D.I. and Zoubak, S.

TITLE Methods and compositions for inducing cell-mediated immune

responses

Patient: WO 0240059-A 7 23-MAY-2002;

American Foundation for Biological Research Inc. (US) ; Mincheff,

Milcho S. (US) ; Loukinov, Dmitri I. (US) ; Zoubak, Serguei (US)

FEATURES

Location/Qualifiers

1..3678

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

7..3678

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD48804.1"

/db\_xref="GI:23386422"

/translation="MRLPASPEHLDMLRHLYQGVVQGNLELYLPTNASLSPLQD

IQEVGYVLIAHNVQVPLQRLIRVGTQQLFEDNYALAVLDNGDPLNNTTPTVGASP

GGRLQLRLSITELKGVLTQRNPOLCYQDTILWKDIFHKNNQALTLIDTNRSRAC

HPSPCKGSRGWSGSDCSLTTRTCAGGCAKCPGLPTCCHEQCAAGCTGPKHS

DCLALHPNHSICELCHPALVYNTDTFESNPPEGRYTFGASCVTACPVNYLSTDV

GCSLTLCPLHNQEVTAEDGTORCEKCSKPCRVCYGLGNEHLREVRAVTSANIQEFAG

CKKIFGSIALFLPESDPASNTAPLQEQVFTLEITGLTGLYLSAWPSLPSLV

FQNLQVIRGLTHNGAYSLTLQGLISWLGRSLRGLSGGLALIHNTHLGCFVHTVPW

DQLFRNPHQALLHTANRDEDCVGEGLACHQLCARGHCWGPGTQCVCNCQFVRGQEC

VECKRVLOGLPREVYNARHCLPCHPECPQNGSVTCFGEPAQOCVACAHYKDPPECA

RCPSGVKPDLSYMPILWKFPEDEGACOPCINCTHSCVDLDKDGCPAQRASPLTSIV

AVGILLVVLGVVFGILLIKRQOKIRKYTHRRLLQETELVEPLTPSCAMPNQAMRI

LKETELRRKVLGSGAGTVYKGIWIPDGENVKIPVAKIKRENTSPKANEILDEAY

VMAGVSPYVSLGLICLTSTVQLTQMLPCLDHLVRENRGLRGLDGLNMCQIA

KGMSYLEDLVRLVHRDLAARNVLVSPNHVKITDFGLRLDITDETEYHADGKVPKIK

MALESILRRRTTHOSDVWSYGVYVWELTFGAKYGPAREIPDLLLEKGRLPPOPI

CTIDVYMTMKWMDISCRPFRELVSEFSRMAPDPQRFVYVIONEDLGASPLDSTF

YRSLEDDMDGLVDABEYLVPPQGFCCPDAPGAGGVMVHRHRSSTSRGSGDITIG

LEPSEEAAPRLAPSEAGSDVFDGLMGAAKGLQLSLDPLHSPLOQRYSEDPVPL

PSFTDGYVAPLTCSQPQYVNPQVPPQPPREGPLFAARPAAGATLBRAKTLSPGKN

GVWKDVFAFGAVENPYLTPQGGAAQPHPPPAFSPAFDNLVYWDQPPPERGAPPST

BASE COUNT 746 a 1137 c 1091 g 704 t  
ORIGIN

FKGTPTAENPEYILGLDVPV"

Alignment Scores:  
Pred. No.: 0.0666 Length: 3678  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 6 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x AX505114 (1-3678)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 2869 GTGGTCATCCAGAAATGAGGACTTGGGCCCGACCGAGTCCCTTGGACAGACACCTTC 2922

RESULT 8

AR034479  
LOCUS AR034479 3768 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 1 from patent US 5869445.  
ACCESSION AR034479  
VERSION AR034479.1 GI:5950084  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3768)  
AUTHORS Cheever,M.A. and Disis,M.L.

TITLE Methods for eliciting or enhancing reactivity to HER-2/neu protein  
JOURNAL Patent: US 5869445-A 1 09-FEB-1999;  
FEATURES Location/Qualifiers  
1..3768

BASE COUNT 759 a 1171 c 1119 g 719 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.0683 Length: 3768  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 6 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x AR034479 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGACCGAGTCCCTTGGACAGACACCTTC 3012

RESULT 9

AX060704  
LOCUS AX060704 3768 bp DNA linear PAT 22-JAN-2001  
DEFINITION Sequence 2 from Patent WO0100244.  
ACCESSION AX060704  
VERSION AX060704.1 GI:12406101  
KEYWORDS  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Erickson,S. and Schwall,R.

TITLE Methods of treatment using anti-erbB antibody-maytansinoid  
conjugates  
JOURNAL Patent: WO 0100244-A 2 04-JAN-2001;  
Genentech, Inc. (US)

FEATURES Location/Qualifiers  
1..3768

/organism="Homo sapiens"  
/mol\_type="genomic DNA"

BASE COUNT 758 a 1170 c 1121 g 719 t  
ORIGIN

/db\_xref="taxon:9606"

Alignment Scores:  
Pred. No.: 0.0683 Length: 3768  
Score: 83.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 51.88% Indels: 0  
DB: 6 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x AX060704 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGACCGAGTCCCTTGGACAGACACCTTC 3012

RESULT 10

AX201817  
LOCUS AX201817 3768 bp DNA linear PAT 30-AUG-2001  
DEFINITION Sequence 1 from Patent WO0153463.  
ACCESSION AX201817  
VERSION AX201817.1 GI:15391666  
KEYWORDS  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Cheever,M.A. and Hand-Zimmermann,S.  
TITLE Compounds and methods for prevention and treatment of her-2/ neu  
associated malignancies  
JOURNAL Patent: WO 0153463-A 1 26-JUL-2001;  
CORIXA CORPORATION (US)

FEATURES Location/Qualifiers  
1..3768

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

1..3768  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="C5391667"  
/db\_xref="GI:15391667"  
/translation="MELAALCRWGLLLALLPGCAASTQVCTGTDMLRLPASPETHLD  
MLRLYGGQVQVGNLELYLPNASISLFQIQEVOGYVLIAHNQVQVPLQRLRV  
RGQLFEDNALAVLDNGDPLNNTPTVGASPGGLRELQRLSLTEILKGGVLIQRNPQ  
LCYQDTILMKDIFHKNNQLALTLDITNRACHPSPMCKSGSCWSESDCSLTNT  
VCAGGACRCKGPLPTDCHEQCAAGTGPKHSDCLACLFHNSGICELHPALVYNT  
DTFESMPNPGRYTFGACSVTACPNVLTQVGSCTLVCPLNHOEVTAEQGTQCEKC  
SKPCARVCYGLGMEHLREVRVTSANIOEACGCKIFGSLAFIPESFDCGDPASNTAPL  
QPEQLQVFTLEITGYLXISAMPDLSPLSVFQNLQVIFGRILHNGAYSLTQGLGI  
SWIGLSRLRELGSGLALIHNTLHNFVHTVPQDLFRNPHQALLHTANRPEDSCVGG  
LACHQVLCARHGCGWPGTQCVNCSQFLRQECVEECVQLQGLPREYVNAHRLCPHPE  
CQPNQSVTCFGEADQVACAHYKDPFCVACRCPGVKPDLSYMWIMVCKWIDSCRRPFRFL  
PCPINTGSCVDLDDKGCAPAEORASPLTSTISAVVGLLVVLVGVFGILIKRQOKI  
RKYTMRLLIQETELVEPLTPSGAMPQAOAMRLKETELKVKVLGSGAGFYVKGNI  
PDGENVKVPAIKVRENTSPKANKELDEAYVMAGVGSFVSRLLGICLTSTVQLVT  
QLMPYGCILLDHRENRLGSLQDLNMCWQIAKMSYLEDVRLVLRHVDLAARNVLSKP  
NHVITDFGLARLLDIEYHADGKGVPIKMMALLESILRRFTHOSVDMSYGVYWE  
LMTFGAKYDGIIPARETDLLEKGERLPQPICTIDVIMVIMVCKWIDSCRRPFRFL  
VSFESRMARDPQRFVIONEDLGPASPLDSTFVRSLEDDMDGLDABEYLVPOQGF  
FCDDPAPGAGGMVHHRSSSTRSGGDLTLGLEPSEEAAPRSLAPSEGASGSDVFG  
DUGMAAKQLSLPTHDPSPLQRISEDTVPVLPSTGDYVAPLTCSPPQPIYVNPQDVR  
PQPPSPREGPLPAARPAAGATLERPKTLSPKNGVGVKDVAFAGAVENPEYLTTPQGGAA  
PQPHPPPAFSPAFDNLVYWDQDPPERGAPPSTFKGTPTAENPEYILGLDVPV"

BASE COUNT 759 a 1171 c 1119 g 719 t  
ORIGIN

Alignment Scores:

Pred. No.: 0.0683

Score: 83.00

Length: 3768

Matches: 16



JUNC\_SEQ8\_SEQ4 (1-30) x AX384604 (1-3768)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 3012

RESULT 13  
 AX465456  
 LOCUS  
 DEFINITION Sequence 1 from Patent WO0213847.  
 ACCESSION AX465456  
 VERSION AX465456.1 GI:21899819  
 KEYWORDS Homo sapiens (human)  
 SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 Gaiger, A., Cheever, M.A. and Hand-zimmermann, S.  
 Methods for diagnosis and therapy of hematological and  
 virus-associated malignancies  
 Patent: WO 0213847-A 1 21-FEB-2002;  
 CORIXA CORPORATION (US)  
 JOURNAL  
 FEATURES  
 source  
 Location/Qualifiers  
 1..3768  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 1..3768  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD42363.1"  
 /db\_xref="GI:21899820"  
 /translation="MELAAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPEHLD  
 MRHLYOGCVVQGNLELYPTNASLFLQDIOEVGYVLIHANQVROPQLRLRV  
 RGTQLFEDNALAVLDNGDPLNTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPQ  
 LCYQDTILMKDIFHNKQALTLIDTNRACHPSPCKSGRCWGSSEDCOSLRT  
 VCAGGACRCGPPLTDCHEQACAGCTGPKHSDCLACLFHNSGICELHCPALVTYNT  
 DTFESMPNPGRTYFGASCTACPYNTLSDVSGCTLVCPHNOEVAEDGTQRCCK  
 SKPCARYGLGMEHLREVAVTSANTQEFAGCKIFGSLAFLPESDGDPAASNTAPL  
 OPEQLOVFETLEETGYLYISAWPDSLPLDSVFONLOVIRGRIILHNGAYSLTLOGLI  
 SWGLSLRELGSGLALIHNTLHCLFVHTVPMDLFRPNHQALLHTANRDECEVGG  
 LACHOLCARHGCGPGTQCNCQSFQRQCEVVEECRVLQGLPREYVNAHCLPCHPE  
 LCPONGSVTCFPGADQVACAHKDPFCVARGPSGKPDLSYMPWKPFDEEGACQ  
 PCPTNCHSCVDLDKGCAPASPLTISAVVGLLVVVLGVFGILIKRQKI  
 RKYTMRLLOETELVEPLTPSGAMPNQAKRIKTELKRVKVLGSGAFGVYKGIWI  
 PDGNNVPAIVKLVRENTSPKANKEILDEAYVMAGVSPVSRLLGICLTSTVOLVT  
 QLMPEYGCLLDHVRENRGLSQDLNLMCQIAKMSYLEVDRLVHRLAARNLVKSP  
 NHVKITDFGLARLLDIDETEHADGGKVPINWMALESILRRRFTHQSDVWSYGVYWE  
 LMTFGAKPIGIPAREIPDLLEKGERLPPICTIDVIMVKNWCMIDSCRPFRREL  
 VSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLDDMDGLVDAEYLVPOQFG  
 FCPDPAAGAGVHHRHRSSTRSGGDLTLGLEPSEEEAPRSLAPSEAGSVDFDG  
 DLGMAAKGLQSLTPHDPSPLORYSEDPTVPLPSETDGYVAPLTCSPQPEVYNQDVR  
 PQPPSPREGPLPAARPAAGATLERPKTILSPCKNGVVKDVFAGGAVENPEYLTPOGGAA  
 PQHPHPPAFSAFDNLYYWDQDPERGAPSTFKGTPTAENPEYLGLDVVP"

BASE COUNT 759 a 1171 c 1119 g 719 t  
 ORIGIN

Alignment Scores:

Pred. No.: 0.0683 Length: 3768  
 Score: 83.00 Matches: 16  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 51.88% Indels: 0  
 DB: 6 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x AX465456 (1-3768)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 3012

RESULT 14

AX467229

LOCUS  
 DEFINITION Sequence 3 from Patent WO0234287.  
 ACCESSION AX467229  
 VERSION AX467229.1 GI:21900511  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 Beier, A.M., Gautam, A. and Mouritsen, S.R.  
 Novel therapeutic vaccine formulations  
 Patent: WO 0234287-A 3 02-MAY-2002;  
 Pharmexa A/S (DK)  
 JOURNAL  
 FEATURES  
 source  
 Location/Qualifiers  
 1..3768  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 1..3768  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD42452.1"  
 /db\_xref="GI:21900512"  
 /translation="MELAAALCRWGLLLALLPPGAASQVCTGTDMLRLPASPEHLD  
 MRHLYOGCVVQGNLELYPTNASLFLQDIOEVGYVLIHANQVROPQLRLRV  
 RGTQLFEDNALAVLDNGDPLNTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPQ  
 LCYQDTILMKDIFHNKQALTLIDTNRACHPSPCKSGRCWGSSEDCOSLRT  
 VCAGGACRCGPPLTDCHEQACAGCTGPKHSDCLACLFHNSGICELHCPALVTYNT  
 DTFESMPNPGRTYFGASCTACPYNTLSDVSGCTLVCPHNOEVAEDGTQRCCK  
 SKPCARYGLGMEHLREVAVTSANTQEFAGCKIFGSLAFLPESDGDPAASNTAPL  
 OPEQLOVFETLEETGYLYISAWPDSLPLDSVFONLOVIRGRIILHNGAYSLTLOGLI  
 SWGLSLRELGSGLALIHNTLHCLFVHTVPMDLFRPNHQALLHTANRDECEVGG  
 LACHOLCARHGCGPGTQCNCQSFQRQCEVVEECRVLQGLPREYVNAHCLPCHPE  
 LCPONGSVTCFPGADQVACAHKDPFCVARGPSGKPDLSYMPWKPFDEEGACQ  
 PCPTNCHSCVDLDKGCAPASPLTISAVVGLLVVVLGVFGILIKRQKI  
 RKYTMRLLOETELVEPLTPSGAMPNQAKRIKTELKRVKVLGSGAFGVYKGIWI  
 PDGNNVPAIVKLVRENTSPKANKEILDEAYVMAGVSPVSRLLGICLTSTVOLVT  
 QLMPEYGCLLDHVRENRGLSQDLNLMCQIAKMSYLEVDRLVHRLAARNLVKSP  
 NHVKITDFGLARLLDIDETEHADGGKVPINWMALESILRRRFTHQSDVWSYGVYWE  
 LMTFGAKPIGIPAREIPDLLEKGERLPPICTIDVIMVKNWCMIDSCRPFRREL  
 VSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLDDMDGLVDAEYLVPOQFG  
 FCPDPAAGAGVHHRHRSSTRSGGDLTLGLEPSEEEAPRSLAPSEAGSVDFDG  
 DLGMAAKGLQSLTPHDPSPLORYSEDPTVPLPSETDGYVAPLTCSPQPEVYNQDVR  
 PQPPSPREGPLPAARPAAGATLERPKTILSPCKNGVVKDVFAGGAVENPEYLTPOGGAA  
 PQHPHPPAFSAFDNLYYWDQDPERGAPSTFKGTPTAENPEYLGLDVVP"

BASE COUNT 758 a 1170 c 1121 g 719 t  
 ORIGIN

Alignment Scores:

Pred. No.: 0.0683 Length: 3768  
 Score: 83.00 Matches: 16  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 51.88% Indels: 0  
 DB: 6 Gaps: 0

JUNC\_SEQ8\_SEQ4 (1-30) x AX467229 (1-3768)

Oy 13 ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 Db 2959 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 3012

RESULT 15

AX481438  
 LOCUS  
 DEFINITION Sequence 52 from Patent WO02055693.  
 ACCESSION AX481438  
 VERSION AX481438.1 GI:22316352  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Kreutzer, R., Limmer, S., Rost, S. and Hadwiger, P.  
Method for inhibiting the expression of a target gene  
Patent: WO 02055693-A 52 18-JUL-2002;  
JOURNAL Ribopharma AG (DE)

FEATURES

source  
1..3768  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 758 a 1170 c 1121 g 719 t  
ORIGIN

Alignment Scores:

Pred. No.:	0.0683	Length:	3768
Score:	83.00	Matches:	16
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	2
Query Match:	51.88%	Indels:	0
DB:	6	Gaps:	0

JUNC\_SEQ8\_SEQ4 (1-30) x AX481438 (1-3768)

QY	13	ValThrPheGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe	30
Db	2959	GTGGTCATCCAGAAATGAGGACTTGGCCCGCCAGCCACTCCCTTGGACAGCACCTTC	3012

Search completed: October 16, 2003, 11:01:46  
Job time : 436.208 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.  
OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 234.794 seconds  
(without alignments)  
3105.423 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ2RES991\_  
Perfect score: 161  
Sequence: 1 ERGCPAEQRASPTVFQEDLGSPSPMDSTF 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delep 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framet\_p2n\_model -DEV=xlp  
-Q/cgn2\_1/USPTO.spool/P/HOLLERAN480/runat\_15102003\_131913\_20548/app\_query.fasta\_1.4685  
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 @CGN\_1\_1\_12645 @runat\_15102003\_131913\_20548 -NCPU=3  
-NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pln:\*
- 20: em\_gss\_vit:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_man:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*

29: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	84	52.2	422	10	BF554233	BF554233 UI-R-C0-h
2	84	52.2	524	13	BQ828479	BQ828479 LL6in1002
3	84	52.2	527	14	CB719751	CB719751 AMGNNUC:N
4	84	52.2	568	13	BQ831943	BQ831943 LL6in1122
5	84	52.2	577	13	BQ829915	BQ829915 LL6in2239
6	84	52.2	592	10	BE382065	BE382065 601272225
7	84	52.2	629	10	BE374596	BE374596 601225384
8	84	52.2	666	13	BU671195	BU671195 NISC_Lr06
9	84	52.2	715	13	BU052728	BU052728 UI-M-EX0-
10	84	52.2	726	12	BI648082	BI648082 603279290
11	84	52.2	770	12	BI645946	BI645946 603274921
12	84	52.2	773	12	BI157032	BI157032 602921306
13	84	52.2	778	12	BG863431	BG863431 602977093
14	84	52.2	785	12	BI688650	BI688650 603311262
15	84	52.2	787	12	BI156730	BI156730 602922150
16	84	52.2	844	12	BI158626	BI158626 602922395
17	84	52.2	881	12	BI645770	BI645770 603275203
18	84	52.2	887	14	CA980253	CA980253 AGENCOURT
19	84	52.2	984	13	BU523863	BU523863 AGENCOURT
20	84	52.2	1988	11	BC023725	BC023725 Mus muscu
21	84	52.2	3110	11	AK031542	AK031542 Mus muscu
22	84	52.2	3372	11	BC046553	BC046553 Mus muscu
23	84	52.2	4323	11	AK031099	AK031099 Mus muscu
24	84	52.2	4463	11	AK083669	AK083669 Mus muscu
25	81.5	50.6	618	28	AZ699313	AZ699313 RPECI-23-2
26	81	50.3	220	9	AW763060	AW763060 ur60c11.y
27	81	50.3	483	4	BX530001	BX530001 R2PD Mus
28	81	50.3	552	10	BF022141	BF022141 uv47c11.y
29	81	50.3	654	12	BI648646	BI648646 603277119
30	81	50.3	655	12	BI557778	BI557778 603236957
31	81	50.3	695	12	BI557170	BI557170 603239144
32	81	50.3	718	12	BI108229	BI108229 602902263
33	81	50.3	749	12	BI554781	BI554781 603236758
34	81	50.3	754	12	BI651822	BI651822 603298821
35	81	50.3	771	12	BI154609	BI154609 602904778
36	81	50.3	891	10	BF101641	BF101641 601753564
37	78	48.4	278	10	BF923905	BF923905 QV4-NT025
38	78	48.4	294	10	BF757857	BF757857 CM4-CT057
39	78	48.4	337	10	BE091682	BE091682 IL2-BT073
40	78	48.4	348	9	AI909847	AI909847 QV-BT225-
41	78	48.4	378	14	CB266642	CB266642 1005398 H
42	78	48.4	423	12	BM703962	BM703962 UI-E-CK1-
43	78	48.4	464	2	HSM067221	Bx477782 Homo sapi
44	78	48.4	496	12	BM795191	BM795191 K-EST0076
45	78	48.4	505	12	BM854421	BM854421 K-EST0136

ALIGNMENTS

RESULT 1  
BF554233  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BF554233 422 bp mRNA linear EST 12-DEC-2000  
UI-R-C0-hp-c-01-0-UI.r1 UI-R-C0 Rattus norvegicus cDNA clone  
UI-R-C0-hp-c-01-0-UI 5', mRNA sequence.  
BF554233  
BF554233.1 GI:11663963  
EST  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 422)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548  
 COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@iowa.edu  
 cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
 clones will be available through Research Genetics (www.resgen.com)  
 This clone is also available through the I.M.A.G.E. Consortium at  
 LLNL (info@image.llnl.gov). IMAGE ID= 1781731  
 Seq primer: M13 Forward.

# FEATURES

source  
 1..422  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-C0-hp-c-01-0-UI"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="UI-R-C0"  
 /notes="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; The UI-R-C0 library is a subtracted library derived from the UI-R-A1 and UI-R-E1 libraries. The UI-R-A1 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, and muscle. The UI-R-E1 library consisted of a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C0) was constructed as follows: PCR amplified cDNA inserts from a pool of UI-R-A1 and UI-R-E1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the pooled UI-R-A1 and UI-R-E1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C0 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)."

BASE COUNT 90 a 115 c 125 g 92 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.0938 Length: 422  
 Score: 84.00 Matches: 16  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 52.17% Indels: 0  
 DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x BF554233 (1-422)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30  
 ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 122 GTGGTCATCCAGAACGAGGACTTGGGCCCATCCAGCCCATGACAGTACCTTC 175  
 ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 2  
 BQ828479

LOCUS BQ828479 524 bp mRNA linear EST 15-SEP-2002  
 DEFINITION LL6in100257 AFT024-subtracted library Mus musculus cDNA 5' similar to Neu/ERBB-2 p185, mRNA sequence.  
 ACCESSION BQ828479  
 VERSION BQ828479.1 GI:22860534  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 524)  
 AUTHORS Hackney,J.A., Charbord,P., Brunk,B.P., Stoeckert,C.J., Lemischka,I.R. and Moore,K.A.  
 TITLE A molecular profile of a hematopoietic stem cell niche  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (20), 13061-13066 (2002)  
 MEDLINE 22247628  
 PUBMED 12226475  
 COMMENT Contact: Moore, Kateri A.  
 Department of Molecular Biology  
 Princeton University  
 217 Lewis Thomas Laboratory, Washington Road, Princeton, NJ 08544, USA  
 Tel: 609 258 0605  
 Fax: 609 258 2759  
 Email: kmoore@molbio.princeton.edu  
 These ESTs are derived from a subtracted cDNA library enriched for gene products expressed by a hematopoietic stem cell-supporting stromal cell line, AFT024.  
 Seq primer: M13Reverse or T7.

# FEATURES

source  
 1..524  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /tissue\_type="Fetal Liver"  
 /cell\_type="Stromal cell"  
 /cell\_line="AFT024"  
 /dev\_stage="Embryonic day 14-14.5"  
 /lab\_host="DH10B"  
 /clone\_lib="AFT024-subtracted library"  
 /notes="Organ: Fetal Liver; Vector: Sport 1; Site.1: Sal I; Site.2: Not I; Two directionally cloned cDNA libraries were made from a hematopoietic stem cell-supporting stromal cell line (AFT024) and from a non-supporting stromal cell line (2018). Subtractive hybridization was performed by hybridization of the target, AFT024, single stranded cDNA library in psport1 to biotinylated RNA transcribed from the driver, 2018 cDNA library in psport2 with inserts cloned in the complementary orientation. The AFT024-subtracted library contains 4.2x105 clones and is depleted of common housekeeping gene products eg. beta-actin and enriched for transcripts specific to AFT024. For detailed protocols and additional information please see our website at http://stromalcell.princeton.edu."

BASE COUNT 110 a 149 c 157 g 107 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.124 Length: 524  
 Score: 84.00 Matches: 16  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 52.17% Indels: 0  
 DB: 13 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x BQ828479 (1-524)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30  
 ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 270 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCATGACAGTACCTTC 323  
 ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||

```

RESULT 3
CB719751
LOCUS
DEFINITION AMGNNUC:NRDGI-00186-C7-A nrdg1 (10855) Rattus norvegicus cDNA clone
ACCESSION nrdg1-00186-c7 5', mRNA sequence.
VERSION CB719751
KEYWORDS CB719751.1 GI:29776893
SOURCE EST.
ORGANISM Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 527)
AUTHORS Angen EST Program
TITLE Angen Rat EST Program
JOURNAL Unpublished
COMMENT Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00186 row: c column: 7.
FEATURES
Location/Qualifiers
source
1..527
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="nrdg1-00186-c7"
/tissue_type="Dorsal Root Ganglia"
/clone_lib="nrdg1 (10855)"
/note="vector: pSPORT1; Site_1: SalI; Site_2: NotI; rat
dorsal root ganglia"
BASE COUNT 105 a 159 c 156 g 107 t
ORIGIN
Alignment Scores:
Pred. No.: 0.125 Length: 527
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 14 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x CB719751 (1-527)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 87 GTGTCATCCAGAACGAGGACTTGGCCCATCCAGCCCATGGACAGTACCTTC 140
||| |||||||||||||||||||||||||||||||||||||||||||||||||||
RESULT 4
BQ831943
LOCUS
DEFINITION LL6in11220T7 AFT024-subtracted library Mus musculus cDNA 5' similar
to Neu/ERBB-2 pl85, mRNA sequence.
ACCESSION BQ831943
VERSION BQ831943.1 GI:22863998
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 568)
AUTHORS Hackney,J.A., Charbord,P., Brunk,B.P., Stoeckert,C.J., Lemischka
,I.R. and Moore,K.A.
TITLE A molecular profile of a hematopoietic stem cell niche
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (20), 13061-13066 (2002)
MEDLINE 22247628
PUBMED 12226475
COMMENT Contact: Moore, Kateri A.
Department of Molecular Biology
Princeton University
217 Lewis Thomas Laboratory, Washington Road, Princeton, NJ 08544,
USA

Tel: 609 258 0605
Fax: 609 258 2759
Email: knoore@molbio.princeton.edu
These ESTs are derived from a subtracted cDNA library enriched for
gene products expressed by a hematopoietic stem cell-supporting
stromal cell line, AFT024.
Seq primer: M13Reverse or T7.
Location/Qualifiers
source
1..568
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57Bl/6J"
/db_xref="taxon:10090"
/tissue_type="Fetal Liver"
/cell_type="Stromal cell"
/cell_line="AFT024"
/dev_stage="Embryonic day 14-14.5"
/lab_host="DH10B"
/clone_lib="AFT024-subtracted library"
/note="Organ: Fetal Liver; Vector: Sport 1; Site_1: Sal I;
Site_2: Not I; Two directionally cloned cDNA libraries
were made from a hematopoietic stem cell-supporting
stromal cell line (AFT024) and from a non-supporting
stromal cell line (2018). Subtractive hybridization was
performed by hybridization of the target, AFT024, single
stranded cDNA library in pSport1 to biotinylated RNA
transcribed from the driver, 2018 cDNA library in pSport2
with inserts cloned in the complementary orientation. The
AFT024-subtracted library contains 4.2x105 clones and is
depleted of common housekeeping gene products eg.
beta-actin and enriched for transcripts specific to
AFT024. For detailed protocols and additional information
please see our website at
http://stromalcell.princeton.edu."
BASE COUNT 122 a 173 c 159 g 112 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 0.138 Length: 568
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 13 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x BQ831943 (1-568)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 156 GTGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCATGGACAGCCTTC 209
||| |||||||||||||||||||||||||||||||||||||||||||||||||||
RESULT 5
BQ829915
LOCUS
DEFINITION LL6in22393 AFT024-subtracted library Mus musculus cDNA 5' similar
to Neu/ERBB-2 pl85, mRNA sequence.
ACCESSION BQ829915
VERSION BQ829915.1 GI:22861970
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 577)
AUTHORS Hackney,J.A., Charbord,P., Brunk,B.P., Stoeckert,C.J., Lemischka
,I.R. and Moore,K.A.
TITLE A molecular profile of a hematopoietic stem cell niche
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (20), 13061-13066 (2002)
MEDLINE 22247628
PUBMED 12226475
COMMENT Contact: Moore, Kateri A.
Department of Molecular Biology
Princeton University

```

217 Lewis Thomas Laboratory, Washington Road, Princeton, NJ 08544,  
USA  
Tel: 609 258 0605  
Fax: 609 258 2759  
Email: kmoore@molbio.princeton.edu  
These ESTs are derived from a subtracted cDNA library enriched for  
gene products expressed by a hematopoietic stem cell-supporting  
stromal cell line, AFT024.  
Seq primer: M13reverse or T7.

## FEATURES

## source

1. .577  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57Bl/6J"  
/db\_xref="taxon:10090"  
/tissue\_type="Fetal Liver"  
/cell\_type="Stromal cell"  
/cell\_line="AFT024"  
/dev\_stage="Embryonic day 14-14.5"  
/lab\_host="DH10B"  
/clone\_lib="AFT024-subtracted library"  
/note="Organ: Fetal Liver; Vector: Sport 1; Site\_1: Sal I;  
Site\_2: Not I; Two directionally cloned cDNA libraries  
were made from a hematopoietic stem cell-supporting  
stromal cell line (AFT024) and from a non-supporting  
stromal cell line (2018). Subtractive hybridization was  
performed by hybridization of the target, AFT024, single  
stranded cDNA library in pSport1 to biotinylated RNA  
transcribed from the driver, 2018 cDNA library in pSport2  
with inserts cloned in the complementary orientation. The  
AFT024-subtracted library contains 4.2x105 clones and is  
depleted of common housekeeping gene products eg.  
beta-actin and enriched for transcripts specific to  
AFT024. For detailed protocols and additional information  
please see our website at  
http://stromalcell.princeton.edu."

BASE COUNT 123 a 173 c 170 g 110 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.141 Length: 577  
Score: 84.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 52.17% Indels: 0  
DB: 13 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x B0829915 (1-577)  
QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30  
Db 167 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCATGGACGACCTTC 220

RESULT 6  
BE382065 592 bp mRNA linear EST 21-JUL-2000  
LOCUS 601272225F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:3601422 5',  
DEFINITION mRNA sequence.  
ACCESSION BE382065.1 GI:9327430  
VERSION EST.  
KEYWORDS Mus musculus (house mouse)  
SOURCE National Institutes of Health, Mammalian Gene Collection (MGC)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 592)  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
AUTHORS Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.

BE382065 592 bp mRNA linear EST 21-JUL-2000  
LOCUS 601272225F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:3601422 5',  
DEFINITION mRNA sequence.  
ACCESSION BE382065.1 GI:9327430  
VERSION EST.  
KEYWORDS Mus musculus (house mouse)  
SOURCE National Institutes of Health, Mammalian Gene Collection (MGC)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 592)  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
AUTHORS Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.

BE382065 592 bp mRNA linear EST 21-JUL-2000  
LOCUS 601272225F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:3601422 5',  
DEFINITION mRNA sequence.  
ACCESSION BE382065.1 GI:9327430  
VERSION EST.  
KEYWORDS Mus musculus (house mouse)  
SOURCE National Institutes of Health, Mammalian Gene Collection (MGC)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 592)  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
AUTHORS Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.

BE382065 592 bp mRNA linear EST 21-JUL-2000  
LOCUS 601272225F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:3601422 5',  
DEFINITION mRNA sequence.  
ACCESSION BE382065.1 GI:9327430  
VERSION EST.  
KEYWORDS Mus musculus (house mouse)  
SOURCE National Institutes of Health, Mammalian Gene Collection (MGC)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 592)  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
AUTHORS Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM8787 row: a column: 07  
High quality sequence stop: 575.

## FEATURES

## source

1. .592  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3601422"  
/tissue\_type="tumor, biopsy sample"  
/dev\_stage="10 months, virgin"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Mam1"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH"

BASE COUNT 116 a 194 c 169 g 113 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.145 Length: 592  
Score: 84.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 52.17% Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x BE382065 (1-592)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30  
Db 76 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCATGGACGACCTTC 129

## RESULT 7

BE374596 629 bp mRNA linear EST 21-JUL-2000  
LOCUS 601225384F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:3583556 5',  
DEFINITION mRNA sequence.  
ACCESSION BE374596  
VERSION BE374596.1 GI:9319959  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

BE374596 629 bp mRNA linear EST 21-JUL-2000  
LOCUS 601225384F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:3583556 5',  
DEFINITION mRNA sequence.  
ACCESSION BE374596  
VERSION BE374596.1 GI:9319959  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 629)  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
AUTHORS Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM8740 row: h column: 21  
High quality sequence stop: 580.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

1. .629  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3583556"

/tissue\_type="tumor, biopsy sample"  
 /dev\_stage="10 months, virgin"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP\_Maml"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"

BASE COUNT 118 a 204 c 181 g 126 t

ORIGIN

#### Alignment Scores:

Pred. No.: 0.157 Length: 629  
 Score: 84.00 Matches: 16  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 52.17% Indels: 0  
 DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x BE374596 (1-629)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerPrometAspSerThrphe 30

Db 8 GTGGTCATCCAGAACGAGGACTTAGCCCTCCAGCCCCATGGACAGCACCTTC 61

#### RESULT 8

BU671195

LOCUS

DEFINITION NISC\_Lr06b08.y1 NCI\_CGAP\_Pr49 Rattus norvegicus cDNA clone

IMAGE:5598279 5', mRNA sequence.

ACCESSION BU671195

VERSION BU671195.1

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 666)

AUTHORS NCI-CGAP

TITLE http://www.ncbi.nlm.nih.gov/ncicgap.

JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

COMMENT Tumor Gene Index

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

CDNA Library Preparation:

CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL

Sequencing Center (NISC)

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

info@image.llnl.gov

Plate: L1AM12385 Row: C Column: 16

Seq primer: M13Rp1 reverse primer (ABI).

Location/Qualifiers

1..666

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

/clone="IMAGE:5598279"

/sex="male"

/tissue\_type="ventral prostate, pool of 3-, 5-, and 7-days

post-castration"

/dev\_stage="adult, 11 week"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NCI\_CGAP\_Pr49"

/note="Organ: prostate; Vector: pCMV-SPORT6.ccdB; Site\_1:

NotI; Site\_2: EcoRV; Cloned unidirectionally. Primer:

Oligo dT. Pool of 3 primary libraries: NCI\_CGAP\_Pr30

(ventral prostate from 11 wk male, 3 days post-castration

, average insert size 2 kb), NCI\_CGAP\_Pr40 (ventral

prostate from 11 wk male, 5 days post-castration, average

insert size 1.6 kb) and NCI\_CGAP\_Pr41 (ventral prostate

from 11 wk male, 7 days post-castration, average insert  
 size 2.5 kb). Constructed by Life  
 Technologies/Invitrogen. Note: this is a NCI\_CGAP  
 Library."

BASE COUNT 149 a 165 c 197 g 155 t

ORIGIN

#### Alignment Scores:

Pred. No.: 0.17 Length: 666  
 Score: 84.00 Matches: 16  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 52.17% Indels: 0  
 DB: 13 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x BU671195 (1-666)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerPrometAspSerThrphe 30

Db 563 GTGGTCATCCAGAACGAGGACTTGGGCCATCCAGCCCCATGGACAGTACCTTC 616

#### RESULT 9

BU052728

LOCUS

DEFINITION UI-M-EX0-b-yk-b-08-0-UI.r1 NIH\_BMAP\_EX0 Mus musculus cDNA clone

IMAGE:5719207 5', mRNA sequence.

ACCESSION BU052728

VERSION BU052728.1

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa

CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..715

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6"

/db\_xref="taxon:10090"

/clone="IMAGE:5719207"

/tissue\_type="whole brain"

/dev\_stage="embryo 15.5 dpc"

/lab\_host="DH10B (T1 phage resistant)"

/clone\_lib="NIH\_BMAP\_EX0"

/note="Organ: brain; Vector: pYX-Asc; Site\_1: EcoR I;

Site\_2: Not I; The library was constructed according to

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured mRNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with an

oligo-dT primer containing a Not I site. Double stranded

cDNA was size selected according to mRNA size fraction,

ligated with EcoR I adaptor, digested with Not I, and then

cloned directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

, is GTGGCGTGAA. This library was created for the

University of Iowa Mouse Brain Molecular Anatomy Project

(BMAP): 'Gene Discovery in the Developing Mouse Nervous

System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

```

BASE COUNT      146 a      216 c      209 g      144 t
ORIGIN

Alignment Scores:
Pred. No.:      0.186      Length:      715
Score:          84.00      Matches:    16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:    52.17%      Indels:    0
DB:            13      Gaps:      0

JUNC_SEQ8_SEQ2RES991_ (1-30) x BU052728 (1-715)

QY      13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db      246 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCCATGGACGACCTTC 299

RESULT 10
LOCUS   BI648082
DEFINITION 726 bp mRNA linear EST 12-SEP-2001
MUSCULUS cdna clone IMAGE:5319502 5',
mRNA sequence.
ACCESSION BI648082
VERSION   BI648082.1 GI:15562318
KEYWORDS EST.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 726)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished
COMMENT  Contact: Robert Strausberg, Ph.D.
        Email: cgabbs-r@mail.nih.gov
        Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
        cDNA Library Preparation: Life Technologies, Inc.
        CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
        DNA Sequencing by: Incyte Genomics, Inc.
        Clone distribution: MGC clone distribution information can be
        found through the I.M.A.G.E. Consortium/LLNL at:
        http://image.llnl.gov
        Plate: LLAM1810 row: c column: 23
        High quality sequence stop: 724.
        Location/Qualifiers
            1..726
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="129, C57BL/6J, FVB/N"
                /db_xref="taxon:10090"
                /clone="IMAGE:5319502"
                /tissue_type="tumor, gross tissue"
                /dev_stage="10 months"
                /lab_host="DH10B"
                /clone_lib="NCI_CGAP_Mam3"
                /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt;
                Site_2: NotI; Cloned unidirectionally. Primer: Oligo df.
                Library constructed by Life Technologies. Investigators
                providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
                Reference for transgenic model: Xu et al., Nature Genetics
                22, 37-43 (1999)."
                22, 219 c      210 g      148 t

BASE COUNT      149 a      219 c      210 g      148 t
ORIGIN

Alignment Scores:
Pred. No.:      0.19      Length:      726
Score:          84.00      Matches:    16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:    52.17%      Indels:    0
DB:            12      Gaps:      0

JUNC_SEQ8_SEQ2RES991_ (1-30) x BI645946 (1-770)

QY      13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db      219 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCCATGGACGACCTTC 272

RESULT 12
LOCUS   BI645946
DEFINITION 773 bp mRNA linear EST 05-JUL-2001
MUSCULUS cdna clone IMAGE:5061649 5',
mRNA sequence.
ACCESSION BI645946
VERSION   BI645946.1 GI:15560182
KEYWORDS EST.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 770)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished
COMMENT  Contact: Robert Strausberg, Ph.D.
        Email: cgabbs-r@mail.nih.gov
        Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
        cDNA Library Preparation: Life Technologies, Inc.
        CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
        DNA Sequencing by: Incyte Genomics, Inc.
        Clone distribution: MGC clone distribution information can be
        found through the I.M.A.G.E. Consortium/LLNL at:
        http://image.llnl.gov
        Plate: LLAM1799 row: i column: 06
        High quality sequence stop: 770.
        Location/Qualifiers
            1..770
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="129, C57BL/6J, FVB/N"
                /db_xref="taxon:10090"
                /clone="IMAGE:5315405"
                /tissue_type="tumor, gross tissue"
                /dev_stage="10 months"
                /lab_host="DH10B"
                /clone_lib="NCI_CGAP_Mam3"
                /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt;
                Site_2: NotI; Cloned unidirectionally. Primer: Oligo df.
                Library constructed by Life Technologies. Investigators
                providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
                Reference for transgenic model: Xu et al., Nature Genetics
                22, 37-43 (1999)."
                22, 245 c      214 g      154 t

BASE COUNT      157 a      245 c      214 g      154 t
ORIGIN

Alignment Scores:
Pred. No.:      0.205      Length:      770
Score:          84.00      Matches:    16
Percent Similarity: 88.89%      Conservative: 0
Best Local Similarity: 88.89%      Mismatches: 2
Query Match:    52.17%      Indels:    0
DB:            12      Gaps:      0

JUNC_SEQ8_SEQ2RES991_ (1-30) x BI645946 (1-770)

QY      13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db      219 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCCATGGACGACCTTC 272

RESULT 12
LOCUS   BI645946
DEFINITION 773 bp mRNA linear EST 05-JUL-2001
MUSCULUS cdna clone IMAGE:5061649 5',
mRNA sequence.
ACCESSION BI645946
VERSION   BI645946.1 GI:15560182
KEYWORDS EST.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 770)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL  Unpublished
COMMENT  Contact: Robert Strausberg, Ph.D.
        Email: cgabbs-r@mail.nih.gov
        Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
        cDNA Library Preparation: Life Technologies, Inc.
        CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
        DNA Sequencing by: Incyte Genomics, Inc.
        Clone distribution: MGC clone distribution information can be
        found through the I.M.A.G.E. Consortium/LLNL at:
        http://image.llnl.gov
        Plate: LLAM1799 row: i column: 06
        High quality sequence stop: 770.
        Location/Qualifiers
            1..770
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="129, C57BL/6J, FVB/N"
                /db_xref="taxon:10090"
                /clone="IMAGE:5315405"
                /tissue_type="tumor, gross tissue"
                /dev_stage="10 months"
                /lab_host="DH10B"
                /clone_lib="NCI_CGAP_Mam3"
                /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt;
                Site_2: NotI; Cloned unidirectionally. Primer: Oligo df.
                Library constructed by Life Technologies. Investigators
                providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
                Reference for transgenic model: Xu et al., Nature Genetics
                22, 37-43 (1999)."
                22, 245 c      214 g      154 t

BASE COUNT      157 a      245 c      214 g      154 t
ORIGIN

```



```

ACCESSION      BI157032
VERSION        BI157032.1  GI:14617033
KEYWORDS       EST.
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 773)
AUTHORS        NIH-MGC http://mgc.nci.nih.gov/
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
               cDNA Library Preparation: Life Technologies, Inc.
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: LLAM1166 row: 1 column: 02
               High quality sequence stop: 772.
               Location/Qualifiers
               1..773
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="129 C57BL/6J.FVB/N"
               /db_xref="taxon:10090"
               /clone="IMAGE:5061649"
               /tissue_type="tumor, gross tissue"
               /dev_stage="10 months"
               /lab_host="DH10B"
               /clone_lib="NCI_CGAP_Mam3"
               /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
               Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
               Library constructed by Life Technologies. Investigators
               providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
               Reference for transgenic model: Xu et al., Nature Genetics
               22, 37-43 (1999)."
```

BASE COUNT 157 a 239 c 219 g 158 t

ORIGIN

Alignment Scores:

Pred. No.:	0.206	Length:	773
Score:	84.00	Matches:	16
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	2
Query Match:	52.17%	Indels:	0
DB:	12	Gaps:	0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x BI157032 (1-773)

```

QY      13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
       111
       226 GTGGTCATCCAGACGAGGACTAGGCCCTCCAGCCCATGGACGACCTTC 279
       Db

RESULT 13
LOCUS   BG863431
DEFINITION 602797093F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4918006 5',
          mRNA sequence.
ACCESSION BG863431
VERSION   BG863431.1  GI:14213969
KEYWORDS  EST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 778)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
```

```

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10830 row: j column: 23
High quality sequence start: 4
High quality sequence stop: 600.
               Location/Qualifiers
               1..778
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="NMRI"
               /db_xref="taxon:10090"
               /clone="IMAGE:4918006"
               /tissue_type="tumor, gross tissue"
               /dev_stage="5 months"
               /lab_host="DH10B"
               /clone_lib="NCI_CGAP_Mam4"
               /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
               Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
               Library constructed by Life Technologies. Investigators
               providing samples: Lothar Hennighausen/Priscilla Furth,
               NIH Reference for transgenic model: Li et al., Cell Growth
               and Differentiation 7, 3-11 (1996)."
```

BASE COUNT 173 a 219 c 225 g 161 t

ORIGIN

Alignment Scores:

Pred. No.:	0.207	Length:	778
Score:	84.00	Matches:	16
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	2
Query Match:	52.17%	Indels:	0
DB:	12	Gaps:	0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x BG863431 (1-778)

```

QY      13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
       111
       431 GTGGTCATCCAGACGAGGACTAGGCCCTCCAGCCCATGGACGACCTTC 484
       Db

RESULT 14
LOCUS   BI688650
DEFINITION 603311262F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5347334 5',
          mRNA sequence.
ACCESSION BI688650
VERSION   BI688650.1  GI:15651279
KEYWORDS  EST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 785)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: Jeffrey Green M.D.
               cDNA Library Preparation: Life Technologies, Inc.
               DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: LLAM11882 row: k column: 15
               High quality sequence stop: 615.
```

FEATURES  
source

Location/Qualifiers  
1. .785  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5347334"  
/sex="female, virgin"  
/tissue\_type="infiltrating ductal carcinoma"  
/dev\_stage="5 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Mam6"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Library constructed by Life Technologies. Investigator  
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 156 a 265 c 214 g 150 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.21 Length: 785  
Score: 84.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 52.17% Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x BI688650 (1-785)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30  
Db 100 GTGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCATGGACACACCTTC 153

## RESULT 15

LOCUS BI156730 787 bp mRNA linear EST 05-JUL-2001  
DEFINITION 602922150F1 NCI\_CGAP\_Mam3 Mus musculus cDNA clone IMAGE:5062237 5',  
mRNA sequence.  
ACCESSION BI156730  
VERSION BI156730.1 GI:14616731  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 787)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM1168 row: d column: 14  
High quality sequence stop: 748.

FEATURES  
source

Location/Qualifiers  
1. .787  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="129/C57BL/6J,FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5062237"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="10 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI\_CGAP\_Mam3"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH  
Reference for transgenic model: Xu et al., Nature Genetics  
22, 37-43 (1999)."

BASE COUNT 166 a 253 c 214 g 154 t  
ORIGIN

## Alignment Scores:

Pred. No.: 0.211 Length: 787  
Score: 84.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 52.17% Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x BI156730 (1-787)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30  
Db 68 GTGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCATGGACACACCTTC 121

Search completed: October 16, 2003, 17:04:04  
Job time : 235.794 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: October 16, 2003, 11:08:07 : Search time 31.0111 Seconds  
(without alignments)  
2540.503 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ2RES991\_  
Perfect score: 161  
Sequence: 1 ERGCPAQRASPVTFQNEGLGPPSPMDSTF 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPto\_spool\_p/HOLLERAN480/runat\_15102003\_131915\_20662/app\_query.fasta\_1.4685  
-DB=PublishedApplications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-NAXLEN=2000000000 -USER=HOLLERAN480 -SCGN\_1\_1\_397 @curat\_15102003\_131915\_20662  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PTCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PTCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	84	52.2	3771	10 US-09-854-356-11	Sequence 11, Appl

2	84	52.2	3955	10 US-09-870-759-117	Sequence 117, App
3	84	52.2	3955	10 US-09-854-356-10	Sequence 10, Appl
4	84	52.2	3955	12 US-09-751-708A-117	Sequence 117, App
5	78	48.4	1115	14 US-10-102-806-165	Sequence 165, App
6	78	48.4	1713	12 US-10-378-393-14	Sequence 14, Appl
7	78	48.4	1755	10 US-09-930-125-6	Sequence 6, Appl
8	78	48.4	1767	10 US-09-930-125-7	Sequence 7, Appl
9	78	48.4	1773	10 US-09-930-125-4	Sequence 4, Appl
10	78	48.4	1806	10 US-09-930-125-5	Sequence 5, Appl
11	78	48.4	2411	12 US-10-378-393-10	Sequence 10, Appl
12	78	48.4	3765	12 US-10-207-498-5	Sequence 5, Appl
13	78	48.4	3768	9 US-09-811-123-8	Sequence 8, Appl
14	78	48.4	3768	9 US-09-811-115-2	Sequence 2, Appl
15	78	48.4	3768	10 US-09-854-356-9	Sequence 9, Appl
16	78	48.4	3768	10 US-09-930-125-1	Sequence 1, Appl
17	78	48.4	3768	12 US-10-313-644-1	Sequence 1, Appl
18	78	48.4	4473	11 US-09-441-411-5	Sequence 5, Appl
19	78	48.4	4473	12 US-10-101-510-81	Sequence 81, Appl
20	78	48.4	4473	14 US-10-146-473-32	Sequence 32, Appl
21	78	48.4	4473	14 US-10-207-655-44	Sequence 44, Appl
22	78	48.4	4530	10 US-09-877-177-11	Sequence 11, Appl
23	78	48.4	4530	12 US-10-007-926A-119	Sequence 119, App
24	78	48.4	4530	12 US-10-101-510-124	Sequence 124, App
25	78	48.4	4530	12 US-10-338-730-1	Sequence 1, Appl
26	78	48.4	4530	14 US-10-177-293-125	Sequence 125, App
27	78	48.4	4543	10 US-09-769-508-1	Sequence 1, Appl
28	78	48.4	4606	12 US-09-971-392-70	Sequence 70, Appl
29	78	48.4	4642	14 US-10-198-846-10896	Sequence 10896, A
30	78	48.4	9274	9 US-09-811-123-7	Sequence 7, Appl
31	78	48.4	9274	9 US-09-811-115-1	Sequence 1, Appl
32	66	41.0	201	13 US-10-109-213-3	Sequence 3, Appl
33	60	37.3	23432	9 US-09-764-869-1332	Sequence 1332, Ap
34	60	37.3	23432	14 US-10-091-504-1332	Sequence 25, Appl
35	59	36.6	1938	12 US-10-422-264-25	Sequence 27, Appl
36	59	36.6	2316	12 US-10-422-264-27	Sequence 23, Appl
37	59	36.6	2604	12 US-09-862-658-1	Sequence 22, Appl
38	59	36.6	3320	10 US-10-175-696-22	Sequence 29, Appl
39	59	36.6	3320	14 US-10-422-264-29	Sequence 29, Appl
40	59	36.6	3384	12 US-09-318-995-28706	Sequence 28706, A
41	57.5	35.7	363	11 US-09-967-768A-292	Sequence 292, App
42	57.5	35.7	2691	13 US-10-028-748-1	Sequence 1, Appl
43	57.5	35.7	2691	13 US-10-002-600-66	Sequence 66, Appl
44	57.5	35.7	2785	13 US-10-027-632-114433	Sequence 114433,
45	57	35.4	602	13	

ALIGNMENTS

RESULT 1  
US-09-854-356-11  
; Sequence 11, Application US/09854356  
; Patent No. US20020177567A1  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Gheysen, Dirk  
; APPLICANT: Corixa Corporation  
; APPLICANT: SmithKline Beecham Biologicals S. A.  
; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
; FILE REFERENCE: 014058-009810PC  
; CURRENT APPLICATION NUMBER: US/09/854,356  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: US 09/493,480  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/117,976  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 3771  
; TYPE: DNA  
; ORGANISM: Mus sp.  
; FEATURE:  
; OTHER INFORMATION: mouse HER-2/neu

US-09-854-356-11

Alignment Scores:

Pred. No.: 0.00383 Length: 3771  
Score: 84.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 52.17% Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x US-09-854-356-11 (1-3771)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30

Db 2962 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCCCATGGACAGCACCTTC 3015

RESULT 2

US-09-870-759-117

; Sequence 117, Application US/09870759

; Patent No. US20020177551A1

; GENERAL INFORMATION:

; APPLICANT: TERNAN, David S

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE

; FILE REFERENCE: 870759

; CURRENT APPLICATION NUMBER: US/09/870,759

; PRIOR FILING DATE: 2002-01-14

; PRIOR APPLICATION NUMBER: US 60/208,128

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 117

; LENGTH: 3955

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (17)..(3799)

; OTHER INFORMATION:

US-09-870-759-117

Alignment Scores:

Pred. No.: 0.00403 Length: 3955  
Score: 84.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 52.17% Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x US-09-870-759-117 (1-3955)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30

Db 2990 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCCCATGGACAGCACCTTC 3043

RESULT 3

US-09-854-356-10

; Sequence 10, Application US/09854356

; Patent No. US20020177567A1

; GENERAL INFORMATION:

; APPLICANT: Cheever, Martin A.

; APPLICANT: Gheysen, Dirk

; APPLICANT: Corixa Corporation

; APPLICANT: Smithline Beecham Biologicals S. A.

; TITLE OF INVENTION: HER-2/neu Fusion Proteins

; FILE REFERENCE: 014058-009810PC

; CURRENT APPLICATION NUMBER: US/09/854,356

; CURRENT FILING DATE: 2001-05-09

; PRIOR APPLICATION NUMBER: US 09/493,480

; PRIOR FILING DATE: 2000-01-28

; PRIOR APPLICATION NUMBER: US 60/117,976

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 10  
; LENGTH: 3955  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: rat HER-2/neu cDNA  
; NAME/KEY: CDS  
; LOCATION: (26)..(3799)  
; OTHER INFORMATION: rat HER-2/neu protein  
; NAME/KEY: misc.feature  
; LOCATION: (26)..(1990)  
; OTHER INFORMATION: extracellular domain (ECD) of rat HER-2/neu  
; NAME/KEY: misc.feature  
; LOCATION: (2057)..(3796)  
; OTHER INFORMATION: intracellular domain (ICD) of rat HER-2/neu  
; NAME/KEY: misc.feature  
; LOCATION: (2188)..(3022)  
; OTHER INFORMATION: kinase domain (KD) of rat HER-2/neu  
; NAME/KEY: misc.feature  
; LOCATION: (2999)..(3796)  
; OTHER INFORMATION: phosphorylation domain (PD) of rat HER-2/neu  
; NAME/KEY: misc.feature  
; LOCATION: (2999)..(3173)  
; OTHER INFORMATION: preferred portion of the phosphorylation domain  
; OTHER INFORMATION: (delta PD) of rat HER-2/neu  
US-09-854-356-10

Alignment Scores:

Pred. No.: 0.00403 Length: 3955  
Score: 84.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 52.17% Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x US-09-854-356-10 (1-3955)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30

Db 2990 GTGGTCATCCAGAACGAGGACTTAGGCCCTCCAGCCCCCATGGACAGCACCTTC 3043

RESULT 4

US-09-751-708A-117

; Sequence 117, Application US/09751708A

; Publication No. US20030157113A1

; GENERAL INFORMATION:

; APPLICANT: TERNAN, David S

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE

; FILE REFERENCE: 751708

; CURRENT APPLICATION NUMBER: US/09/751,708A

; CURRENT FILING DATE: 2002-10-15

; PRIOR APPLICATION NUMBER: US 60/173,371

; PRIOR FILING DATE: 1999-12-28

; NUMBER OF SEQ ID NOS: 166

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 117

; LENGTH: 3955

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (17)..(3799)

; OTHER INFORMATION:

US-09-751-708A-117

Alignment Scores:

Pred. No.: 0.00403 Length: 3955  
Score: 84.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 52.17% Indels: 0  
DB: 12 Gaps: 0

```
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-751-708A-117 (1-3955)
Oy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 2990 GTGGTCATCCAGAACAGGAGCTTGGGCCCATCCAGCCCGCCAGTACCTTC 3043
RESULT 5
US-10-102-806-165
; Sequence 165, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL03P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 165
; LENGTH: 1115
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (390)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-102-806-165
Alignment Scores:
Pred. No.: 0.0085 Length: 1115
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 14 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-10-102-806-165 (1-1115)
Oy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 446 GTGGTCATCCAGAACAGGAGCTTGGGCCCATCCAGCCCGCCAGTACCTTC 499
RESULT 6
US-10-378-393-14
; Sequence 14, Application US/10378393
; Publication No. US20030182668A1
; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVATED
; FILE REFERENCE: D0254 NP
; CURRENT APPLICATION NUMBER: US/10/378,393
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,889
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 1713
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-378-393-14
Alignment Scores:
Pred. No.: 0.0136 Length: 1713
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 12 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-10-378-393-14 (1-1713)
Oy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 904 GTGGTCATCCAGAACAGGAGCTTGGGCCCATCCAGCCCGCCAGTACCTTC 957
RESULT 7
US-09-930-125-6
; Sequence 6, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-6
Alignment Scores:
Pred. No.: 0.014 Length: 1755
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 10 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-930-125-6 (1-1755)
Oy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 937 GTGGTCATCCAGAACAGGAGCTTGGGCCCATCCAGCCCGCCAGTACCTTC 990
RESULT 8
US-09-930-125-4
; Sequence 4, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
```

```
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-4
Alignment Scores:
Pred. No.: 0.0141 Length: 1767
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 0 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-930-125-4 (1-1767)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 937 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 990
RESULT 9
US-09-930-125-7
; Sequence 7, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-7
Alignment Scores:
Pred. No.: 0.0141 Length: 1773
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 0 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-930-125-7 (1-1773)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 961 GTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 1014
RESULT 10
US-09-930-125-5
; Sequence 5, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
```

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1806
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-930-125-5
Alignment Scores:
Pred. No.: 0.0144 Length: 1806
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 10 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-930-125-5 (1-1806)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 994 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 1047
RESULT 11
US-10-378-393-10
; Sequence 10, Application US/10378393
; Publication No. US2003018268A1
; GENERAL INFORMATION:
; APPLICANT: Bol, David K.
; APPLICANT: Carboni, Joan M.
; APPLICANT: Rowley, Ronald B.
; APPLICANT: Wong, Tai W.
; APPLICANT: Lee, Francis Y.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMALS EXPRESSING CONSTITUTIVELY ACTIVAT
; FILE REFERENCE: D0254 NP
; CURRENT APPLICATION NUMBER: US/10/378,393
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/360,889
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-378-393-10
Alignment Scores:
Pred. No.: 0.0198 Length: 2411
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 12 Gaps: 0
JUNC_SEQ8_SEQ2RES991_ (1-30) x US-10-378-393-10 (1-2411)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 1602 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 1655
RESULT 12
US-10-207-498-5
; Sequence 5, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
```

```
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HERGULIN AND HER3
; FILE REFERENCE: 30448.103-US-01
; CURRENT APPLICATION NUMBER: US/10/207,498
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-207-498-5

Alignment Scores:
Pred. No.: 0.0323      Length: 3765
Score: 78.00          Matches: 14
Percent Similarity: 88.89%      Conservative: 2
Best Local Similarity: 77.78%    Mismatches: 2
Query Match: 48.45%            Indels: 0
DB: 12                     Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-10-207-498-5 (1-3765)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3012

RESULT 13
US-09-811-123-8
; Sequence 8, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
; TITLE OF INVENTION: ANTIBODY-MAYTANSINOID CONJUGATES
; FILE REFERENCE: GENENT.073A2
; CURRENT APPLICATION NUMBER: US/09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-123-8

Alignment Scores:
Pred. No.: 0.0323      Length: 3768
Score: 78.00          Matches: 14
Percent Similarity: 88.89%      Conservative: 2
Best Local Similarity: 77.78%    Mismatches: 2
Query Match: 48.45%            Indels: 0
DB: 9                     Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-811-123-8 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3012
```

```
RESULT 14
US-09-811-115-2
; Sequence 2, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-115-2

Alignment Scores:
Pred. No.: 0.0323      Length: 3768
Score: 78.00          Matches: 14
Percent Similarity: 88.89%      Conservative: 2
Best Local Similarity: 77.78%    Mismatches: 2
Query Match: 48.45%            Indels: 0
DB: 9                     Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-811-115-2 (1-3768)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGTCCCTTGGACAGCACCTTC 3012

RESULT 15
US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: Intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc_feature
```

; LOCATION: (2968)..(3144)  
; OTHER INFORMATION: preferred portion of the phosphorylation domain  
; OTHER INFORMATION: (delta PD) of human HER-2/neu  
US-09-854-356-9

Alignment Scores:  
Pred. No.: 0.0323 Length: 3768  
Score: 78.00 Matches: 14  
Percent Similarity: 88.89% Conservative: 2  
Best Local Similarity: 77.78% Mismatches: 2  
Query Match: 48.45% Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x US-09-854-356-9 (1-3768)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30  
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCCAGCCAGTCCTTGGACACACCTTC 3012

Search completed: October 17, 2003, 03:53:07  
Job time : 35.0111 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 22:55:19 ; Search time 7.60488 Seconds  
(without alignments)  
1741.185 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ2RES991\_

Perfect score: 161

Sequence: 1 ERGCPAEQRASPVTFQNEGLGSPSPMDSTF 30

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPRO\_spool\_p/HOLLERAN480/runat\_15102003\_131913\_20570/app\_query.fasta\_1.4685  
-DB=Issued\_Patents\_NA -QFWT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480.ecgn\_1.1.403.@runat\_15102003\_131913\_20570 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents\_NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCFUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	84	52.2	3955	1	US-08-229-515A-14
2	84	52.2	3955	1	US-08-645-865-14
3	78	48.4	3768	2	US-08-625-101-1
4	78	48.4	3768	2	US-08-356-786-1
5	78	48.4	4473	2	US-09-048-804-1
6	78	48.4	4473	3	US-09-056-105-26
7	78	48.4	4530	1	US-08-229-515A-9
8	78	48.4	4530	1	US-08-645-865-9
9	78	48.4	4530	4	US-09-167-322-4
10	78	48.4	4530	4	US-09-527-487-1
11	78	48.4	4530	4	US-09-877-177A-11
12	68	42.2	2385	2	US-09-146-283-3

13	68	42.2	2385	3	US-08-579-823A-3	Sequence 3, Appli
14	68	42.2	2385	3	US-09-344-195-3	Sequence 3, Appli
15	66	41.0	153	3	US-08-776-251-3	Sequence 3, Appli
16	66	41.0	201	4	US-09-200-355-3	Sequence 3, Appli
17	66	41.0	816	3	US-08-776-251-10	Sequence 10, Appli
18	59	36.6	1938	4	US-09-547-435-25	Sequence 25, Appli
19	59	36.6	2316	4	US-09-547-435-27	Sequence 27, Appli
20	59	36.6	2604	4	US-09-547-435-23	Sequence 23, Appli
21	59	36.6	3384	4	US-09-547-435-29	Sequence 29, Appli
22	57.5	35.7	2691	4	US-09-020-743-1	Sequence 1, Appli
23	55	34.2	4092	3	US-09-306-595C-5	Sequence 5, Appli
24	55	34.2	4092	4	US-09-925-388-5	Sequence 5, Appli
25	54	33.5	960	3	US-08-651-136C-1	Sequence 1, Appli
26	54	33.5	960	4	US-09-229-911A-1	Sequence 1, Appli
27	54	33.5	5046	4	US-09-548-938A-5	Sequence 5, Appli
28	54	33.5	111282	4	US-09-754-250-3	Sequence 3, Appli
29	53.5	33.2	864	3	US-09-022-983-3	Sequence 3, Appli
30	53.5	33.2	2340	3	US-09-022-983-4	Sequence 4, Appli
31	53.5	33.2	2477	3	US-09-490-692-3	Sequence 3, Appli
32	53.5	33.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
33	53.5	33.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli
34	53	32.9	669	4	US-09-252-991A-9807	Sequence 9807, Ap
35	53	32.9	1020	2	US-08-282-137C-1	Sequence 1, Appli
36	53	32.9	1353	4	US-09-252-991A-9985	Sequence 9985, Ap
37	53	32.9	42571	4	US-09-810-347-3	Sequence 3, Appli
38	53	32.9	536165	4	US-09-214-808-1	Sequence 1, Appli
39	52.5	32.6	444	4	US-09-252-991A-15154	Sequence 15154, A
40	52.5	32.6	465	4	US-09-252-991A-14734	Sequence 14734, A
41	52.5	32.6	621	4	US-09-252-991A-15020	Sequence 15020, A
42	52.5	32.6	1107	4	US-09-252-991A-932	Sequence 932, App
43	52.5	32.6	1266	4	US-09-252-991A-14898	Sequence 14898, A
44	52.5	32.6	1527	4	US-09-252-991A-1035	Sequence 1035, Ap
45	52.5	32.6	1531	2	US-08-983-352-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-229-515A-14  
; Sequence 14, Application US/08229515A  
; Patent No. 5518885  
; GENERAL INFORMATION:  
; APPLICANT: RAZIUDIN  
; APPLICANT: SARKAR, FAZIL H  
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN  
; TITLE OF INVENTION: NEOPlastic DISEASE  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NEEDLE & ROSENBERG PC  
; STREET: 127 Peachtree Street, Suite 1200  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: usa  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/229,515A  
; FILING DATE: 19 APR 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PERRYMAN, DAVID G  
; REGISTRATION NUMBER: 33,438  
; REFERENCE/DOCKET NUMBER: 1414.608  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404-688-0770  
; TELEFAX: 404-688-9880  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3955 base pairs

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-229-515A-14

Alignment Scores:
Pred. No.: 0.00361 Length: 3955
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-08-229-515A-14 (1-3955)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
Db 2990 GTGGTCATCCAGAACGAGGACTTGGGCCCATCCAGCCCCATGACAGTACCTTC 3043

RESULT 2
US-08-645-865-14
; Sequence 14, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; TITLE OF INVENTION: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:
; LENGTH: 3955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-645-865-14

Alignment Scores:
Pred. No.: 0.00361 Length: 3955
Score: 84.00 Matches: 16
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 2
Query Match: 52.17% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-08-645-865-14 (1-3955)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
Db 2990 GTGGTCATCCAGAACGAGGACTTGGGCCCATCCAGCCCCATGACAGTACCTTC 3043
```

```
RESULT 3
US-08-625-101-1
; Sequence 1, Application US/08625101
; Patent No. 5869445
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,101
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3765
US-08-625-101-1

Alignment Scores:
Pred. No.: 0.0285 Length: 3768
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-08-625-101-1 (1-3768)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCATCCAGTACCTTCGACAGACCTTC 3012

RESULT 4
US-08-356-786-1
; Sequence 1, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
```

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3768
; OTHER INFORMATION: /note= "product = "cerB-b2""
;
; US-08-356-786-1
;
; Alignment Scores:
; Pred. No.: 0.0285 Length: 3768
; Score: 78.00 Matches: 14
; Percent Similarity: 88.89% Conservative: 2
; Best Local Similarity: 77.78% Mismatches: 2
; Query Match: 48.45% Indels: 0
; DB: 2 Gaps: 0
;
; JUNC_SEQ8_SEQ2RES991_ (1-30) x US-08-356-786-1 (1-3768)
;
; Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
; ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; Db 2959 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGCTCCCTTGGACAGACCTTC 3012
;
; RESULT 5
; US-09-048-804-1
; Sequence 1, Application US/09048804
; Patent No. 5968748
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748rls LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
;
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/048,804
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legaard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2913
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4473 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; ANTI-SENSE: No
;
; US-09-048-804-1
;
; Alignment Scores:
; Pred. No.: 0.0355 Length: 4473
; Score: 78.00 Matches: 14
; Percent Similarity: 88.89% Conservative: 2
; Best Local Similarity: 77.78% Mismatches: 2
; Query Match: 48.45% Indels: 0
; DB: 2 Gaps: 0
;
; JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-048-804-1 (1-4473)
;
; Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
; ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; Db 3133 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGCTCCCTTGGACAGACCTTC 3186
;
; RESULT 6
; US-09-056-105-26
; Sequence 26, Application US/09056105
; Patent No. 6287569
; GENERAL INFORMATION:
; APPLICANT: KIPPS, THOMAS J.
; APPLICANT: WU, YUNQI
; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
; FILE REFERENCE: 233/221
; CURRENT APPLICATION NUMBER: US/09/056,105
; CURRENT FILING DATE: 1998-04-06
; EARLIER APPLICATION NUMBER: 60/043,467
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
;
; US-09-056-105-26
;
; Alignment Scores:
; Pred. No.: 0.0355 Length: 4473
; Score: 78.00 Matches: 14
; Percent Similarity: 88.89% Conservative: 2
; Best Local Similarity: 77.78% Mismatches: 2
; Query Match: 48.45% Indels: 0
; DB: 3 Gaps: 0
;
; JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-056-105-26 (1-4473)
;
; Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerPrometAspSerThrPhe 30
; ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; Db 3133 GTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCAGCTCCCTTGGACAGACCTTC 3186
;
; RESULT 7
; US-08-229-515A-9
```

```
; Sequence 9, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-229-515A-9

Alignment Scores:
Pred. No.: 0.036 Length: 4530
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-08-229-515A-9 (1-4530)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGTCCCTTGACAGCACCTTC 3162

RESULT 8
US-08-645-865-9
; Sequence 9, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-645-865-9

Alignment Scores:
Pred. No.: 0.036 Length: 4530
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-08-645-865-9 (1-4530)
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAATGAGGACTTGGGCCCGCCAGTCCCTTGACAGCACCTTC 3162

RESULT 9
US-09-167-322-4
; Sequence 4, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
```

```

; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-167-322-4

Alignment Scores:
Pred. No.: 0.036 Length: 4530
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-167-322-4 (1-4530)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 10
US-09-527-487-1
; Sequence 1, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(3915)
US-09-527-487-1

Alignment Scores:
Pred. No.: 0.036 Length: 4530
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-527-487-1 (1-4530)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 11
US-09-877-177A-11
; Sequence 11, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-New Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
US-09-877-177A-11

Alignment Scores:
Pred. No.: 0.036 Length: 4530
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-877-177A-11 (1-4530)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 3109 GTGGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTC 3162

RESULT 12
US-09-146-283-3
; Sequence 3, Application US/09146283
; Patent No. 5976546
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/146,283
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-09-146-283-3

Alignment Scores:
Pred. No.: 0.556 Length: 2385
Score: 68.00 Matches: 15
Percent Similarity: 68.97% Conservative: 5
Best Local Similarity: 51.72% Mismatches: 7
Query Match: 42.24% Indels: 2
DB: 2 Gaps: 1

JUNC_SEQ8_SEQ2RES991_ (1-30) x US-09-146-283-3 (1-2385)

```





**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 28.4941 Seconds  
(without alignments)  
2842.104 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ2RES991\_

Perfect score: 161

Sequence: 1 ERGCPABQRASPVTFQNEGLGPPSPMDSTF 30

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cqn2.1/USPFO\_spool\_p/HOLLERAN480/runat\_15102003\_131912\_20526/app\_query.fasta\_1.4685  
-DB=N\_Geneseq\_19Jun03 -QMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 -CGN\_1.1\_1758 -runat\_15102003\_131912\_20526 -NCFU=6 -ICPU=3  
-NO\_MMAP -LARGUEURY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	153	95.0	2763	24	ABA92252 Mouse Her-2/neu ex
2	153	95.0	2781	21	ABA92253 Her-2/neu extracel
3	84	52.2	3771	21	AAA9737 Mouse Her-2/neu cd
4	84	52.2	3771	22	AAH42210 Nucleotide sequenc
5	84	52.2	3771	24	ABA92251 Mouse Her-2/neu cd
6	84	52.2	3955	16	AAT01590 Rat neu promoter.
7	84	52.2	3955	21	AAA9753 Rat HER-2/neu prot
8	78	48.4	1115	21	AAF21778 Human breast and o
9	78	48.4	1755	24	AAD32746 Human cDNA for the
10	78	48.4	1767	24	AAD32744 Human cDNA for the
11	78	48.4	1773	24	AAD32747 Human cDNA for the
12	78	48.4	1806	24	AAD32745 Human cDNA for the
13	78	48.4	3600	21	AAA89736 Human HER-2/neu co
14	78	48.4	3678	24	ABK86207 cDNA encoding huma
15	78	48.4	3768	17	AAT40739 HER-2/neu oncogene
16	78	48.4	3768	20	AAK01912 Human HER-2/neu on
17	78	48.4	3768	21	AAA09455 Human heregulin 2
18	78	48.4	3768	22	AAH23392 Human HER-2/neu pr
19	78	48.4	3768	24	ABZ35744 Human ERBB2 polynu
20	78	48.4	3768	24	ABX09987 Human ERBB2 DNA fr
21	78	48.4	3768	24	AAD43935 Human HER-2 cDNA.
22	78	48.4	3768	24	AAD43986 Human Her2 antigen
23	78	48.4	3768	24	ABV78168 Human ERBB2 DNA SE
24	78	48.4	3768	24	AAD32743 Human Her-2/neu pr
25	78	48.4	3768	24	ABA92250 Human Her-2/neu CD
26	78	48.4	3768	24	ABK10730 Human Her-2/neu DN
27	78	48.4	3768	24	ABL1709 Human polynucleoti
28	78	48.4	3768	24	ABK14058 Human HER2 (ERBB2)
29	78	48.4	4299	14	AAQ46083 Sequence encoding the
30	78	48.4	4472	21	AAA14812 cDNA encoding the
31	78	48.4	4473	19	ABO76220 Human tumour antig
32	78	48.4	4473	20	AAZ31071 HER-2 nucleic acid
33	78	48.4	4473	24	ABZ34969 Human gene express
34	78	48.4	4473	24	AAD38904 Human Her-2 DNA.
35	78	48.4	4530	16	AAT01585 Human HER2 gene.
36	78	48.4	4530	18	AA71253 Nucleotide sequenc
37	78	48.4	4530	21	AA260815 Human tyrosine kin
38	78	48.4	4530	22	AAD19731 Human gene express
39	78	48.4	4530	24	ABZ35012 Breast carcinoma r
40	78	48.4	4530	24	ABV94128 Human HER2-neu SEQ
41	78	48.4	4530	24	ABN85585 Human cDNA differe
42	78	48.4	4530	24	ABK83918 Breast cancer asso
43	78	48.4	4530	25	ACC03139 Human Her2/Neu enc
44	78	48.4	4530	25	ABQ83856 HER2 transgene pla
45	78	48.4	9274	22	ANF4297

ALIGNMENTS

RESULT 1  
ABA92252  
ID ABA92252 standard; cDNA: 2763 BP.  
XX  
AC ABA92252;  
XX  
XX  
XX 17-JUN-2002 (first entry)  
XX  
DE Mouse Her-2/neu extracellular-phosphorylation domain fusion cDNA.  
XX  
KW Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;  
KW receptor; mouse; gene therapy; gene; ss.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2763

```

FT      /*tag= a
XX      /product= "Her-2/neu ECD-PD fusion"
PN      WO200212341-A2.
XX
PD      14-FEB-2002.
XX
XX      03-AUG-2001; 2001WO-US24283.
XX
XX      03-AUG-2000; 2000US-0632507.
XX
XX      (CORI-) CORIXA CORP.
PA      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX      Cheever MA, Gheysen D;
XX
DR      WPI; 2002-241743/29.
XX      P-PSDB; AAMS1152.
XX
XX      Her-2/neu fusion protein for treating or preventing cancer by eliciting
XX      or enhancing an immune response to the protein, has Her-2/neu
XX      extracellular domain fused to Her-2/neu intracellular or
XX      phosphorylation domain
XX
PS      Disclosure; Fig 23; 141pp; English.
XX
XX      The present sequence is that of cDNA encoding a fusion between
XX      the extracellular domain (ECD) and the phosphorylation domain (PD)
XX      murine Her-2/neu oncoprotein (see AAMS1151). The Her-2/neu gene is
XX      amplified and the oncoprotein is overexpressed in a variety of
XX      human cancers, including breast, ovarian, colon, lung and prostate
XX      cancer. Her-2/neu overexpression correlates with a poor prognosis
XX      in breast and ovarian cancers. The invention provides Her-2/neu
XX      fusion proteins, nucleic acids encoding them, viral vectors, and
XX      vaccines comprising the fusion proteins or nucleic acid molecules.
XX      In preferred fusion proteins, the ECD of Her-2/neu is fused to a
XX      Her-2/neu intracellular domain or PD (or its DeltapB fragment). An
XX      immune response to Her-2/neu protein is elicited or enhanced by
XX      administering the fusion protein in the form of a vaccine, or by
XX      transfecting cells of an animal ex vivo with a nucleic acid
XX      encoding the fusion protein, and delivering the transfected cells
XX      to the animal. The fusion proteins, nucleic acids, and isolated
XX      specific T-cells are useful for inhibiting the development of a
XX      cancer, especially breast, ovarian, colon, lung or prostate cancer
XX      in a patient. T cells that specifically react with a Her-2/neu
XX      fusion protein can be used to remove tumour cells from a sample in
XX      order to inhibit the development of cancer in a patient.
XX
SQ      Sequence 2763 BP; 571 A; 855 C; 772 G; 565 T; 0 other;

Alignment Scores:
Pred. No.:      1.58e-12      Length:      2763
Score:          153.00      Matches:      29
Percent Similarity: 96.67%      Conservative: 0
Best Local Similarity: 96.67%      Mismatches: 1
Query Match:     95.03%      Indels:      0
DB:             24      Gaps:        0

JUNC_SEQ8_SEQ2RES991_ (1-30) x ABA92252 (1-2763)

Qy      1 GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
Dy      1918 GAACGAGGCTGCGCCAGCAGACGAGAGCCAGCCAGTGCAGTCTCAGAACGAGGACTTA 1977
Qy      21 GlyProSerSerProMetAspSerThrPhe 30
Dy      1978 GGCCCTCCAGCCCGCCAGGACGACGACCTTC 2007

RESULT 2
ABA92253
ID      ABA92253 standard; cDNA; 2781 BP.
XX
AC      ABA92253;

17-JUN-2002 (first entry)
Her-2/neu extracellular-phosphorylation domain-TcP0 fusion cDNA.
Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
receptor; TcP0; mouse; gene therapy; gene; ss.
Chimeric - Mus musculus.
Chimeric - Unidentified.
Key      Location/Qualifiers
CDS      1..2781
/*tag= a
/*product= "Her-2/neu ECD-PD-TcP0 fusion"
WO200212341-A2.
14-FEB-2002.
03-AUG-2001; 2001WO-US24283.
03-AUG-2000; 2000US-0632507.
(CORI-) CORIXA CORP.
(SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
Cheever MA, Gheysen D;
WPI; 2002-241743/29.
P-PSDB; AAMS1153.
Her-2/neu fusion protein for treating or preventing cancer by eliciting
or enhancing an immune response to the protein, has Her-2/neu
extracellular domain fused to Her-2/neu intracellular or
phosphorylation domain
Disclosure; Fig 25; 141pp; English.
The present sequence is that of cDNA encoding a fusion between
the extracellular domain (ECD) and the phosphorylation domain (PD)
murine Her-2/neu oncoprotein (see AAMS1151) plus a C-terminal
TcP0 motif that improves immunogenicity. The Her-2/neu gene is
amplified and the oncoprotein is overexpressed in a variety of
human cancers, including breast, ovarian, colon, lung and prostate
cancer. Her-2/neu overexpression correlates with a poor prognosis
in breast and ovarian cancers. The invention provides Her-2/neu
fusion proteins, nucleic acids encoding them, viral vectors, and
vaccines comprising the fusion proteins or nucleic acid molecules.
In preferred fusion proteins, the ECD of Her-2/neu is fused to a
Her-2/neu intracellular domain or PD (or its DeltapB fragment). An
immune response to Her-2/neu protein is elicited or enhanced by
administering the fusion protein in the form of a vaccine, or by
transfecting cells of an animal ex vivo with a nucleic acid
encoding the fusion protein, and delivering the transfected cells
to the animal. The fusion proteins, nucleic acids, and isolated
specific T-cells are useful for inhibiting the development of a
cancer, especially breast, ovarian, colon, lung or prostate cancer
in a patient. T cells that specifically react with a Her-2/neu
fusion protein can be used to remove tumour cells from a sample in
order to inhibit the development of cancer in a patient.
Sequence 2781 BP; 574 A; 859 C; 779 G; 569 T; 0 other;

Alignment Scores:
Pred. No.:      1.59e-12      Length:      2781
Score:          153.00      Matches:      29
Percent Similarity: 96.67%      Conservative: 0
Best Local Similarity: 96.67%      Mismatches: 1
Query Match:     95.03%      Indels:      0
DB:             24      Gaps:        0

JUNC_SEQ8_SEQ2RES991_ (1-30) x ABA92253 (1-2781)

```

D	B		2962 GTGGTCATCAGAACGAGGACTTAGGCCCTCCACAGCCCCCATGGACAGCACCTTC 3015
 RESULT 4			
AAH42210			
ID	AAH42210	standard; cDNA; 3771 BP.	
XX	XX		
AC	AAH42210;		
XX	XX		
DT	17-SEP-2001	(first entry)	
XX	XX		
DE	Nucleotide sequence of a murine neu polypeptide from C57Bl6 mice.		
XX	XX		
KW	Neu polypeptide; Her-2; cancer; breast cancer; T cell expansion;		
KW	vaccine; ss.		
XX	XX		
OS	Mus sp.		
XX	XX		
FH	Key	Location/Qualifiers	
FT	CDS	1..3771	
FT	FT	/*tag= a	
FT	FT	/product= "neu polypeptide"	
PN	WO200148205-A2.		
XX	XX		
PD	05-JUL-2001.		
XX	XX		
PF	29-DEC-2000; 2000WO-US35648.		
XX	XX		
PR	29-DEC-1999; 99US-0474382.		
XX	XX		
PA	(CORI-) CORIXA CORP.		
XX	XX		
PI	Spies AG;		
XX	XX		
DR	WPI; 2001-441718/47.		
DR	P-PSDB; AAG62860.		
XX	XX		
PT	Novel isolated murine homolog of human Her-2/neu useful for inhibiting		
PT	development of cancer, preferably breast cancer in a patient -		
XX	XX		
PS	Claim 4; Fig 1A-B; 53pp; English.		
XX	XX		
CC	The present sequence encodes a murine neu polypeptide. It is a murine		
CC	homologue of human Her-2/neu. Neu polypeptides and polynucleotides are		
CC	useful for inhibiting the development of a cancer, preferably breast		
CC	cancer in a patient. They are also useful for stimulating and/or		
CC	expanding T cells specific for Her-2/neu. They are useful for detecting		
CC	and treating Her-2/neu malignancies. Neu polypeptides are useful		
CC	for producing vaccines.		
XX	XX		
SQ	Sequence 3771 BP; 808 A; 1105 C; 1068 G; 790 T; 0 other;		
 Alignment Scores:			
Pred. No.:	0.0342	Length:	3771
Score:	84.00	Matches:	16
Percent Similarity:	88.89%	Conservative:	0
Best Local Similarity:	88.89%	Mismatches:	2
Query Match:	52.17%	Indels:	0
DB:	22	Gaps:	0
JUNC_SEQ8_SEQ2RES991_	{1-30} x AAH42210 {1-3771}		
QY	13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30		
DB			
2962 GTGGTCATCAGAACGAGGACTTAGGCCCTCCACAGCCCCCATGGACAGCACCTTC 3015			
 RESULT 5			
ABA92251			
ID	ABA92251	standard; cDNA; 3771 BP.	
XX	XX		
AC	ABA92251;		
XX	XX		
DT	17-JUN-2002	(first entry)	



JUNC\_SEQ08\_SEQ02RES991\_ (1-30) x AAA89753 (1-3955)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30  
III |  
Db 2990 GTGGTCATCCAGACGAGACTTGGGCCCATCCAGCCCCATGGACAGTACCTC 3043

RESULT 8  
AAF21778  
ID AAF21778 standard; DNA; 1115 BP.  
XX AC AAF21778;  
XX XX  
XX 27-MAR-2001 (first entry)  
DE Human breast and ovarian cancer associated antigen gene SEQ ID 165.  
XX  
XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;  
KW neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;  
KW antidabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;  
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;  
KW Addison's disease; allergy; autoimmune hemolytic anaemia;  
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;  
KW cardiovascular disorder; wound healing; neurological disease; ds.  
XX  
XX Homo sapiens.  
OS  
XX WO200055173-A1.  
PN  
XX  
PD 21-SEP-2000.  
PF  
XX 08-MAR-2000; 2000WO-US05881.  
XX  
PR 12-MAR-1999; 99US-0124270.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Rosen CA, Ruben SM;  
PI  
XX WPI; 2000-611515/58.  
DR P-PSDB; AAB58875.  
XX  
PT New human breast and ovarian cancer associated gene sequences and the  
PT polypeptides encoded by these genes, useful in the prevention,  
PT treatment and diagnosis of cancer, immune disorders, cardiovascular  
PT disorders and neurological diseases -  
XX Claim 1; Page 604; 1299pp; English.

Sequence AAF21614 - AAF259123 represent DNA sequences encoding human proteins AAB58711 - AAB591128. The DNA and protein sequences are associated with breast and ovarian cancer. Included in the invention are sequences AAF22032 - AAF22040 and AAB59129 which are used in the isolation and characterisation of the DNA and protein sequences of the invention. The breast and ovarian cancer associated DNA, protein, agonist or antagonist sequences exhibit cytostatic; immunosuppressive; neotropic; neuroprotective; antiviral; antiallergic; hepatotropic; antidabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic and cardiant activity. The polynucleotide and protein sequences are used in the diagnosis of cancer, particularly breast and ovarian cancer. The nucleic acid sequences, proteins, agonists and antagonists may also be used in the diagnosis, prevention and treatment of immune disorders e.g. Addison's disease, allergies, autoimmune hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; cardiovascular disorders such as myocardial ischemias; wound healing; neurological diseases such as cerebral anoxia and epilepsy; and infectious diseases.

XX Sequence 1115 BP; 210 A; 336 C; 338 G; 222 T; 9 other;  
SQ

Alignment Scores:  
Pred. No.: 0.0601 Length: 1115

Score: 78.00 Matches: 14  
Percent Similarity: 88.89% Conservative: 2  
Best Local Similarity: 77.78% Mismatches: 2  
Query Match: 48.45% Indels: 0  
DB: 21 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x AAF21778 (1-1115)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerPrometAspSerThrPhe 30  
||| |||||||||||||||||||||:|||||:|||||  
Db 446 GTGGTCATCCAGAATGAGGACTTGGCCAGCCAGTCCCTTGGACAGCACCTTC 499

RESULT 9  
AAD32746  
ID AAD32746 standard; cDNA; 1755 BP.  
XX AAD32746;  
XX  
XX  
XX 01-JUL-2002 (first entry)  
XX Human cDNA for the clone HICD\_native\_coding\_region.  
XX Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;  
XX human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..1755  
XX /\*tag= a  
XX /product= "Human protein encoded by cDNA for the clone  
XX HICD\_native\_coding\_region"  
XX /transl\_except= (pos:1741..1752, aa:Leu-Glu)  
XX /note= "CDS does not include stop codon"  
XX /partial  
XX  
XX WO200214503-A2.  
XX  
XX 21-FEB-2002.  
XX  
XX 14-AUG-2001; 2001WO-US41733.  
XX  
XX 14-AUG-2000; 2000US-225152P.  
XX 28-SEP-2000; 2000US-236428P.  
XX 21-FEB-2001; 2001US-270520P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;  
XX McNeill PD, Vedvick TS;  
XX WPI; 2002-280758/32.  
XX P-PSDB; AAE20483.  
XX  
XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,  
XX prevention and diagnosis of cancer, preferably breast cancer  
XX  
XX Example 5; Page 118-119; 129pp; English.  
XX  
XX The invention relates to an isolated Her-2/Neu polypeptide composition  
XX effective for eliciting an immune response. The invention is useful for  
XX eliciting an immune response in a patient, where the patient is human  
XX leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.  
XX The composition is useful for the therapy and diagnosis of cancer.  
XX Preferably breast cancer, in pharmaceutical compositions, e.g., vaccine  
XX and other compositions for the diagnosis, prevention and treatment of  
XX human malignancies, for stimulating and/or expanding T cells specific for  
XX Her-2/Neu polypeptide and for inhibiting the development of cancer in a  
XX patient. The invention is useful for stimulating a T cell response in a  
XX human patient, as probe or primer for nucleic acid hybridisation, to  
XX selectively form duplex molecules with complementary stretches of the  
XX entire Her-2/Neu gene or gene fragments of interest, to isolate a full  
XX length gene from a suitable library, and to direct expression of a

CC polypeptide in appropriate host cells. The composition is useful in  
CC prophylactic or therapeutic applications and for the treatment of cancer,  
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-  
CC associated malignancies. The invention is useful in gene therapy. The  
CC present sequence is human cDNA for the clone HICD\_native\_coding\_region.  
XX  
SQ Sequence 1755 BP; 376 A; 517 C; 531 G; 331 T; 0 other;

Alignment Scores:  
Pred. No.: 0.104 Length: 1755  
Score: 78.00 Matches: 14  
Percent Similarity: 88.89% Conservative: 2  
Best Local Similarity: 77.78% Mismatches: 2  
Query Match: 48.45% Indels: 0  
DB: 21 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x AAD32746 (1-1755)

QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerPrometAspSerThrPhe 30  
||| |||||||||||||||||||||:|||||:|||||  
Db 937 GTGGTCATCCAGAATGAGGACTTGGCCAGCCAGTCCCTTGGACAGCACCTTC 990

RESULT 10  
AAD32744  
ID AAD32744 standard; cDNA; 1767 BP.  
XX  
XX AAD32744;  
XX  
XX 01-JUL-2002 (first entry)  
XX Human cDNA for the clone HICD\_CT\_His\_coding\_region.  
XX  
XX Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;  
XX human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..1764  
XX /\*tag= a  
XX /product= "Human protein encoded by cDNA for the clone  
XX HICD\_CT\_His\_coding\_region"  
XX  
XX WO200214503-A2.  
XX  
XX 21-FEB-2002.  
XX  
XX 14-AUG-2001; 2001WO-US41733.  
XX  
XX 14-AUG-2000; 2000US-225152P.  
XX 28-SEP-2000; 2000US-236428P.  
XX 21-FEB-2001; 2001US-270520P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;  
XX McNeill PD, Vedvick TS;  
XX WPI; 2002-280758/32.  
XX P-PSDB; AAE20481.  
XX  
XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,  
XX prevention and diagnosis of cancer, preferably breast cancer  
XX  
XX Example 5; Page 117-118; 129pp; English.  
XX  
XX The invention relates to an isolated Her-2/Neu polypeptide composition  
XX effective for eliciting an immune response. The invention is useful for  
XX eliciting an immune response in a patient, where the patient is human  
XX leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.  
XX The composition is useful for the therapy and diagnosis of cancer,  
XX preferably breast cancer, in pharmaceutical compositions, e.g., vaccine  
XX and other compositions for the diagnosis, prevention and treatment of

CC human malignancies, for stimulating and/or expanding T cells specific for  
 CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a  
 CC patient. The invention is useful for stimulating a T cell response in a  
 CC human patient, as probe or primer for nucleic acid hybridisation, to  
 CC selectively form duplex molecules with complementary stretches of the  
 CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full  
 CC length gene from a suitable library, and to direct expression of a  
 CC polypeptide in appropriate host cells. The composition is useful in  
 CC prophylactic or therapeutic applications and for the treatment of cancer,  
 CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-  
 CC associated malignancies. The invention is useful in gene therapy. The  
 CC present sequence is human cDNA for the clone HICD\_CT\_His\_coding\_region.  
 XX  
 SQ Sequence 1767 BP; 381 A; 521 C; 529 G; 336 T; 0 other;

Alignment Scores:  
 Pred. No.: 0.105 Length: 1767  
 Score: 78.00 Matches: 14  
 Percent Similarity: 88.89% Conservative: 2  
 Best Local Similarity: 77.78% Mismatches: 2  
 Query Match: 48.45% Indels: 0  
 DB: 24 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x AAD32744 (1-1767)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30  
 Db 937 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 990

RESULT 11  
 AAD32747

ID AAD32747 standard; cDNA; 1773 BP.

AC AAD32747;

XX 01-JUL-2002 (first entry)

DE Human cDNA for the clone HICD\_in\_ppdm\_coding\_sequence.

KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;  
 KW human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1770

FT /\*tag= a

FT /product= "Human protein encoded by cDNA for the clone

FT HICD\_in\_ppdm\_coding\_sequence"

XX WO200214503-A2.

XX 21-FEB-2002.

XX 14-AUG-2001; 2001WO-US41733.

XX 14-AUG-2000; 2000US-225152P.

XX 28-SEP-2000; 2000US-236428P.

XX 21-FEB-2001; 2001US-270520P.

XX (CORI-) CORIXA CORP.

XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;

PI McNeill PD, Vedvick TS;

XX WPI; 2002-280758/32.

XX P-PSDB; AAE20484.

XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,

XX prevention and diagnosis of cancer, preferably breast cancer

XX Example 5; Page 119; 129pp; English.

XX

CC The invention relates to an isolated Her-2/Neu polypeptide composition  
 CC effective for eliciting an immune response. The invention is useful for  
 CC eliciting an immune response in a patient, where the patient is human  
 CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.  
 CC The composition is useful for the therapy and diagnosis of cancer,  
 CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine  
 CC and other compositions for the diagnosis, prevention and treatment of  
 CC human malignancies, for stimulating and/or expanding T cells specific for  
 CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a  
 CC patient. The invention is useful for stimulating a T cell response in a  
 CC human patient, as probe or primer for nucleic acid hybridisation, to  
 CC selectively form duplex molecules with complementary stretches of the  
 CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full  
 CC length gene from a suitable library, and to direct expression of a  
 CC polypeptide in appropriate host cells. The composition is useful in  
 CC prophylactic or therapeutic applications and for the treatment of cancer,  
 CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-  
 CC associated malignancies. The invention is useful in gene therapy. The  
 CC present sequence is human cDNA for the clone HICD\_in\_ppdm\_coding\_  
 XX sequence.

SQ Sequence 1773 BP; 383 A; 528 C; 530 G; 332 T; 0 other;

Alignment Scores:  
 Pred. No.: 0.105 Length: 1773  
 Score: 78.00 Matches: 14  
 Percent Similarity: 88.89% Conservative: 2  
 Best Local Similarity: 77.78% Mismatches: 2  
 Query Match: 48.45% Indels: 0  
 DB: 24 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x AAD32747 (1-1773)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30  
 Db 961 GTGGTCATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGACCTTC 1014

RESULT 12

AAD32745

ID AAD32745 standard; cDNA; 1806 BP.

XX AAD32745;

XX 01-JUL-2002 (first entry)

DE Human cDNA for the clone HICD\_plus\_8\_HIS.

XX Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;

XX human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1803

FT /\*tag= a

FT /product= "Human protein encoded by cDNA for the clone

FT HICD\_plus\_8\_HIS"

FT /transl\_except= (pos:1543..1545, aa:Pro)

XX WO200214503-A2.

XX 21-FEB-2002.

XX 14-AUG-2001; 2001WO-US41733.

XX 14-AUG-2000; 2000US-225152P.

XX 28-SEP-2000; 2000US-236428P.

XX 21-FEB-2001; 2001US-270520P.

XX (CORI-) CORIXA CORP.

XX Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos MD;

PI McNeill PD, Vedvick TS;

```
XX WPI; 2002-280758/32.
DR P-PSDB; AAE20482.
XX
PT Novel isolated Her-2/Neu polypeptide composition useful for therapy,
PT prevention and diagnosis of cancer, preferably breast cancer -
XX
PS Example 5; Page 118; 129pp; English.
XX
CC The invention relates to an isolated Her-2/Neu polypeptide composition
CC effective for eliciting an immune response. The invention is useful for
CC eliciting an immune response in a patient, where the patient is human
CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
CC The composition is useful for the therapy and diagnosis of cancer,
CC preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
CC and other compositions for the diagnosis, prevention and treatment of
CC human malignancies, for stimulating and/or expanding T cells specific for
CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a
CC patient. The invention is useful for stimulating a T cell response in a
CC human patient, as probe or primer for nucleic acid hybridisation, to
CC selectively form duplex molecules with complementary stretches of the
CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full
CC length gene from a suitable library, and to direct expression of a
CC polypeptide in appropriate host cells. The composition is useful in
CC prophylactic or therapeutic applications and for the treatment of cancer,
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-
CC associated malignancies. The invention is useful in gene therapy. The
CC present sequence is human cDNA for the clone HICD_plus_8_HIS.
XX
SQ Sequence 1806 BP; 391 A; 530 C; 544 G; 341 T; 0 other;

Alignment Scores:
Pred. No.: 0.108 Length: 1806
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 24 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x AAD32745 (1-1806)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 994 GTGGTCATCCAGAATGAGGACTTGGGCCCGAGTCCTTGGACAGCACCTTC 1047

RESULT 13
AAA89736
ID AAA89736 standard; DNA; 3600 BP.
XX
AC AAA89736;
XX
DT 12-JAN-2001 (first entry)
XX
DE Human HER-2/neu coding sequence.
XX
KW Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
KW colon cancer; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3600
FT FT /*tag= a
FT FT /product= "HER-2/neu protein"
XX
PN WO200044899-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-US02164.
XX
PR 29-JAN-1999; 99US-0117976.
```

```
XX (CORI-) CORIXA CORP.
PA (SMIK ) SMITHKLINE BEECHAM.
XX
PI Cheever MA, Gheysen D;
XX
DR WPI; 2000-505976/45.
DR P-PSDB; AAB21198, AAB21208.
XX
PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and
PT prostate cancers -
XX
PS Disclosure; Fig 15; 128pp; English.
XX
CC The present sequence encodes the human HER-2/neu protein. HER-2/neu is
CC a member of the tyrosine kinase family of receptor-like glycoproteins
CC and shows homology to the epidermal growth factor receptor (EGFR). It
CC probably plays a part in cell growth and/or differentiation. The
CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising
CC a HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation
CC domain may be used to treat or prevent cancer by eliciting or
CC enhancing an immune response to the HER-2/neu protein. It may be used
CC to treat malignancies such as breast, ovarian, colon, lung and
CC prostate cancers, and may be used as an antigen to vaccinate against
CC these neoplasias.
XX
SQ Sequence 3600 BP; 723 A; 1108 C; 1075 G; 594 T; 0 other;

Alignment Scores:
Pred. No.: 0.248 Length: 3600
Score: 78.00 Matches: 14
Percent Similarity: 88.89% Conservative: 2
Best Local Similarity: 77.78% Mismatches: 2
Query Match: 48.45% Indels: 0
DB: 21 Gaps: 0

JUNC_SEQ8_SEQ2RES991_ (1-30) x AAA89736 (1-3600)
QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
Db 2959 GTGGTCATCCAGAATGAGGACTTGGGCCCGAGTCCTTGGACAGCACCTTC 3012

RESULT 14
ABK86207
ID ABK86207 standard; cDNA; 3678 BP.
XX
AC ABK86207;
XX
DT 24-SEP-2002 (first entry)
XX
DE cDNA encoding human breast cancer antigen, Her2 variant.
XX
KW Human; Her2; cytostatic; antiviral; immunostimulant;
KW cell-mediated immune response; tumour; breast cancer;
KW virus infection; prostate cancer; colorectal cancer; pancreatic cancer;
KW lymphoma; leukaemia; hepatitis virus; lentivirus; herpesvirus;
KW human immunodeficiency virus; HIV; flavivirus; pestivirus; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 7..3678
FT FT /*tag= a
FT FT /product= "Breast cancer antigen Her2 variant"
XX
PN WO200240059-A2.
XX
PD 23-MAY-2002.
XX
PF 01-NOV-2001; 2001WO-US45626.
XX
PR 01-NOV-2000; 2000US-0704232.
```



```

XX (AMBI-) AMERICAN FOUND BIOLOGICAL RES INC.
PA {MINC/} MINCHEFF M S.
PA {LOUK/} LOUKINOV D I.
PA {ZOUB/} ZOUBAK S.
XX
XX Mincheff MS, Loukinov DI, Zoubak S;
XX
XX WPI; 2002-527524/56.
XX P-PSDB; AAU98923.
XX
XX Inducing a cell-mediated immune response against a target antigen,
XX reducing undesired cells and stimulating presentation of an antigen by
XX a cell, comprises administering a polynucleotide encoding a variant of
XX an antigen -
XX
XX Disclosure; Page 128-134; 146pp; English.
XX
XX The invention relates to a method of inducing a cell-mediated immune
XX response against a cell comprising a target antigen (I) in a subject,
XX treating a subject having undesired cells, for example tumour cells
XX or virally infected cells (C), reducing the number of (C) in a subject,
XX and stimulating presentation of (I) by a cell. This is done by
XX administering a polynucleotide (II) encoding a variant of (I), so that
XX (II) expressed in a cell and cell-mediated immune response is induced.
XX The method can be used to treat prostate cancer, breast cancer,
XX colorectal cancer and pancreatic cancer, as well as lymphomas and
XX leukaemias. The method is also useful in treating chronic viral
XX infections such as those caused by hepatitisviruses, lentiviruses
XX (including human immunodeficiency virus (HIV)), herpesviruses and the
XX flaviviruses and pestiviruses. The present sequence represents the coding
XX sequence of human breast cancer antigen, Her2 variant, used as a target
XX antigen in the method of the invention.
XX
XX Sequence 3678 BP; 746 A; 1137 C; 1091 G; 704 T; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 0.254 Length: 3678
XX Score: 78.00 Matches: 14
XX Percent Similarity: 88.89% Conservative: 2
XX Best Local Similarity: 77.78% Mismatches: 2
XX Query Match: 48.45% Indels: 0
XX DB: 24 Gaps: 0
XX
XX JUNC_SEQ08_SEQ2RES991_(1-30) x ABK86207 (1-3678)
XX
XX QY 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
XX ||| |||||||||||||||||||||||||||||||||||||||||||||||||||
XX Db 2869 GTGGTGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCCCTTGGACACACCTTC 2922
XX
XX RESULT 15
XX AAT40739
XX ID AAT40739 standard; cDNA; 3768 BP.
XX
XX AC AAT40739;
XX
XX XX
XX 01-JAN-1997 (first entry)
XX
XX DE HER-2/neu oncogene.
XX
XX HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;
XX breast cancer; ovary cancer; colon cancer; lung cancer;
XX prostate cancer; genetic immunisation; tumour; vaccine; vector;
XX ss.
XX
XX Homo sapiens.
XX
XX OS
XX
XX XX
XX Location/Qualifiers
XX Key 1..3765
XX CDS /*tag= b
XX FT ./note= "nucleotides 2026-3765 (claim 1) code for
XX FT HER-2/neu intracellular domain"
XX
XX

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:09:13 : Search time 432.026 Seconds  
(without alignments)  
2840.777 Million cell updates/sec

Title: JUNC\_SEQ8\_SEQ2RES991\_  
Perfect score: 161  
Sequence: 1 ERGCPAEQRASPVTFQNEIDLGSPSPMDSTF 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=framer-p2n.model -DEV=xlp  
-O/cgn2\_1/USPTO\_spool\_p/HOLLERAN480/runat\_15102003\_131912\_20536/app\_query.fasta\_1.4685  
-DB=GenEmbl -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=HOLLERAN480 -SCGN 1.1 2724 @runat\_15102003\_131912\_20536 -NCPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg\_hum.\*  
31: em.htg\_inv.\*  
32: em.htg\_other.\*  
33: em.htg\_mus.\*  
34: em.htg\_pln.\*  
35: em.htg\_rtd.\*  
36: em.htg\_mam.\*  
37: em.htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	153	95.0	2763	6	AX380942 Sequence
2	153	95.0	2781	6	AX380944 Sequence
3	84	52.2	2083	10	BC027080 Mus muscu
4	84	52.2	3771	6	AX189662 Sequence
5	84	52.2	3771	6	AX189662 Sequence
6	84	52.2	3955	6	AX380925 Sequence
7	84	52.2	3955	6	AX380924 Sequence
8	84	52.2	3955	6	I21129 Sequence 14
9	84	52.2	3955	6	I59750 Sequence 14
10	84	52.2	4694	10	RNNEUR X03362 Rat mRNA fo
11	84	52.2	4695	10	BC046811 Mus muscu
12	84	52.2	4727	10	BC053078 Mus muscu
13	81.5	50.6	186134	10	AY116182 Rattus no
14	81.5	50.6	214019	2	AL591390 Mouse DNA
15	81	50.3	4062	10	AC064803 Mus muscu
16	78	48.4	1755	6	HAMNEU D16295 Mesocrictu
17	78	48.4	1767	6	AX384609 Sequence
18	78	48.4	1773	6	AX384607 Sequence
19	78	48.4	1806	6	AX384610 Sequence
20	78	48.4	3678	6	AX384608 Sequence
21	78	48.4	3678	6	AX505114 Sequence
22	78	48.4	3768	6	AR034479 Sequence
23	78	48.4	3768	6	AX060704 Sequence
24	78	48.4	3768	6	AX201817 Sequence
25	78	48.4	3768	6	AX380923 Sequence
26	78	48.4	3768	6	AX384604 Sequence
27	78	48.4	3768	6	AX465456 Sequence
28	78	48.4	3768	6	AX467229 Sequence
29	78	48.4	3768	6	AX481438 Sequence
30	78	48.4	3780	4	AB008451 Canis fam
31	78	48.4	4473	6	AR080259 Sequence
32	78	48.4	4473	9	AR167390 Sequence
33	78	48.4	4530	6	X03363 Human c-erb
34	78	48.4	4530	6	AR202597 Sequence
35	78	48.4	4530	6	AR283481 Sequence
36	78	48.4	4530	6	AX282577 Sequence
37	78	48.4	4530	6	AX587649 Sequence
38	78	48.4	4530	6	AX644071 Sequence
39	78	48.4	4530	6	BD005474 Cellular
40	78	48.4	4530	6	I21124 Sequence 9
41	78	48.4	4530	6	I59745 Sequence 9
42	78	48.4	4530	9	HUMHER2A M1730 Human tyros
43	75	46.6	528	10	AF393158 Rattus no
44	75	46.6	239732	2	AC098491 Rattus no
45	74	46.0	13450	9	AB096612 Homo sapi

ALIGNMENTS

RESULT 1

```

AX380942
LOCUS       AX380942                2763 bp    DNA
DEFINITION   Sequence 28 from Patent WO0212341.
ACCESSION   AX380942
VERSION     AX380942.1  GI:19575786
KEYWORDS    synthetic construct
SOURCE      synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Cheever,M.A. and Gheysen,D.
  TITLE     Her-2/neu fusion proteins
  JOURNAL   Patent: WO 0212341-A 28 14-FEB-2002;
            CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES             source
   1..2763
     /organism="synthetic construct"
     /mol_type="genomic DNA"
     /db_xref="taxon:32630"
     /note="mouse ECD-PD fusion protein cDNA"
BASE COUNT  571 a  855 c  772 g  565 t
ORIGIN
Alignment Scores:
Pred. No.:      5,57e-12      Length:      2763
Score:          153.00      Matches:      29
Percent Similarity: 96.67%      Conservative: 0
Best Local Similarity: 96.67%      Mismatches: 1
Query Match:     95.03%      Indels:      0
DB:              6      Gaps:      0

JUNC_SEQ8_SEQ2RES991_ (1-30) x AX380942 (1-2763)

QY       1  GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
          |||||||
Db       1918  GAACGAGGCTGCCAGCAGCAGAGAGCCAGCCAGCTGAGCTCTCAGACGAGGACTTA 1977
          |||||||

QY       21  GlyProSerProMetAspSerThrPhe 30
          |||||||
Db       1978  GGCCCTCCAGCCCGCCATGGACAGCACCTTC 2007
          |||||||

RESULT 2
LOCUS       AX380944                2781 bp    DNA
DEFINITION   Sequence 30 from Patent WO0212341.
ACCESSION   AX380944
VERSION     AX380944.1  GI:19575787
KEYWORDS    synthetic construct
SOURCE      synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Cheever,M.A. and Gheysen,D.
  TITLE     Her-2/neu fusion proteins
  JOURNAL   Patent: WO 0212341-A 30 14-FEB-2002;
            CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES             source
   1..2781
     /organism="synthetic construct"
     /mol_type="genomic DNA"
     /db_xref="taxon:32630"
     /note="mouse ECD-PD-TcP0 fusion protein cDNA"
BASE COUNT  574 a  859 c  779 g  569 t
ORIGIN
Alignment Scores:
Pred. No.:      5,61e-12      Length:      2781
Score:          153.00      Matches:      29
Percent Similarity: 96.67%      Conservative: 0
Best Local Similarity: 96.67%      Mismatches: 1
Query Match:     95.03%      Indels:      0
DB:              6      Gaps:      0

```

```

JUNC_SEQ8_SEQ2RES991_ (1-30) x AX380944 (1-2781)

QY       1  GluArgGlyCysProAlaGluGlnArgAlaSerProValThrPheGlnAsnGluAspLeu 20
          |||||||
Db       1918  GAACGAGGCTGCCAGCAGCAGAGCCAGCCAGCTGAGCTCTCAGACGAGGACTTA 1977
          |||||||

QY       21  GlyProSerProMetAspSerThrPhe 30
          |||||||
Db       1978  GGCCCTCCAGCCCGCCATGGACAGCACCTTC 2007
          |||||||

RESULT 3
LOCUS       BC027080                2083 bp    mRNA
DEFINITION   Mus musculus v-erb-b2 erythroblastic leukemia viral oncogene
            homolog 2, neuro/glioblastoma derived oncogene homolog (avian),
            mRNA (CDNA clone MGC:38648 IMAGE:5356166), complete cds.
ACCESSION   BC027080
VERSION     BC027080.1  GI:20071980
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 2083)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,K.H., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stاپleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Kettaman,M., Madan,A., Rodriguez,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
            Butterfield,J.S., Krzywinski,M.I., Skalska,U., Smalius,D.E.,
            Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
            Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL     22388257
MEDLINE    12477932
PUBMED
REFERENCE   2 (bases 1 to 2083)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-APR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgabs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lounseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAC Plate: 54 Row: c Column: 18
            This clone was selected for full length sequencing because it

```

```
source
1. 1993
/organism="Rattus norvegicus"
```

/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/note="rat Her-2/neu cDNA"  
26..3799  
/note="rat Her-2/neu protein"  
/codon\_start=1  
/protein\_id="CAD28393.1"  
/db\_xref="GI:19575770"  
/translation="MELAAWCVGFLALLPPGIAGTQVCTGTDMKLRLLPASPTHLD  
MLRHYGCGVQGNLETVVPANASISFLQDIOEVGYMLIAHNOVKRVPQLRLRV  
RGTOLFEDKVALVLDNRDQDNVAATPGTPEGLRELQLRLSLTEILKGLVLRGNP  
QLTQDVAWLNKDFRKNQALAPYDIDNRSRACPPCAPACKDNHCWGESPEDCOILTG  
TICFSGCARCKRULPTDCCHEQCAAGCTPKHSDCLACLFHNSGICELCHCPALVTYN  
TDFESMHNPEGRYTFGASCTTCTPYNLSTEVGSCITLVCPPNNQEVTAEDGTCCK  
CSKPCARVGLGMEHLRGARAITSDNNVDFDGCKKIFGSLAFLEPFGDPSSGIAP  
LRPQLQVFTLEITGYLYISAMPDLSRLDSVFQNLRIIRGRILHDGAYSLTLQIGL  
IHSGLRSLVLEGLALIHRAHLCFVHVVPMDLFRNPQHALLHSGNRPEDLCVS  
SGLVNSLCARHGCHGWGPTQVCNCSHFLRGQCEVCRVWKGLPREYVSDKRCLPCH  
PECPQNSSEFCFGEADQCAAHYKDSSECVARCPGSKVCPDLSYMPIWKYPDEGI  
QCPPIINCHSCVDLDERGCPAQRASPTVFIATVEGVLLFLVYVGVGILIKRRRQ  
KIRKYMRLRLQETELVEPLTPSKAMPNQAMRLKELTKRVKVLGSGAGFTVYKGI  
WIPDGNVKIPVAIKVLENTSPKANKEILDEAVVMAGVSPVYSRLIGLITSTVOL  
VTOLMPYGCILLDVRHGRGLGSQDILLMCVQIAKMSYLEDLVLRDLAARNLVK  
SPNHVKTIDGLARLLIDETEYHADGCKVPKIMALESILRRFTQSDVWSTGVTV  
WELTFGAKPTDGPAREIDPLELLEKGERLPPOPICTIDVYIMVWKWIDSECRPRFR  
ELVSEFSRMDRQRFVYIQNEDLGPSPMDSFYRSLLEDODMDGLVDAEYLVPQQ  
GFFSPDPTGTTAHRHRSSTRSGGELTGLLEPSEEGPPRSLAPSEGAGSDVF  
DGLAMGVTLQSLDPLQRISEDTPLPPTDGVVAPLACSPQPEVYNQSE  
VOQDPLTPPEGLPPVPAGATLERKTLSPGKNGVVKDVFAGGAVENEYELVPRG  
TASPPHSPAFSPADNLYWDQNSSEQPPPSFEGTPTAENPEYLGLDVVPV"  
26..1990  
/note="extracellular domain (ECD) of rat Her-2/neu"  
2057..3796  
/note="intracellular domain (ICD) of rat Her-2/neu"  
2188..3022  
/note="kinase domain (KD) of rat Her-2/neu"  
2999..3796  
/note="phosphorylation domain (PD) of rat Her-2/neu"  
2999..3173  
/note="preferred portion of the phosphorylation domain  
(delta PD) of rat Her-2/neu"  
BASE COUNT 842 a 1147 c 1136 g 830 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.0546 Length: 3955  
Score: 84.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 52.17% Indels: 0  
DB: 6 Gaps: 0  
JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x AX380924 (1-3955)  
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30  
Db 2990 GTGTCATCCAGACGAGGACTTGGGCCCATCCAGCCCATGGACAGTACCTTC 3043  
RESULT 7  
LOCUS I21129 3955 bp DNA linear PAT 07-OCT-1996  
DEFINITION Sequence 14 from patent US 5518885.  
ACCESSION I21129  
VERSION I21129.1 GI:1601483  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3955)  
AUTHORS Raziuddin and Sarkar, F.H.  
TITLE ERBB2 promoter binding protein in neoplastic disease  
JOURNAL Nature 319 (6050), 226-230 (1986)  
FEATURES Location/Qualifiers

source  
BASE COUNT 842 a 1147 c 1136 g 830 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.0546 Length: 3955  
Score: 84.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 52.17% Indels: 0  
DB: 6 Gaps: 0  
JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x I21129 (1-3955)  
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30  
Db 2990 GTGTCATCCAGACGAGGACTTGGGCCCATCCAGCCCATGGACAGTACCTTC 3043  
RESULT 8  
LOCUS I59750 3955 bp DNA linear PAT 07-OCT-1997  
DEFINITION Sequence 14 from patent US 5654406.  
ACCESSION I59750  
VERSION I59750.1 GI:2478382  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3955)  
AUTHORS Raziuddin and Sarkar, F.Hoque.  
TITLE Antibody to ERBB2 promoter binding factor  
JOURNAL Patent: US 5654406-A 14 05-AUG-1997;  
FEATURES Location/Qualifiers  
source  
BASE COUNT 842 a 1147 c 1136 g 830 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.0546 Length: 3955  
Score: 84.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 52.17% Indels: 0  
DB: 6 Gaps: 0  
JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x I59750 (1-3955)  
Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30  
Db 2990 GTGTCATCCAGACGAGGACTTGGGCCCATCCAGCCCATGGACAGTACCTTC 3043  
RESULT 9  
LOCUS RNNEUR 3955 bp mRNA linear ROD 30-MAR-1995  
DEFINITION Rat mRNA for neu oncogene (p185) encoding an epidermal growth  
factor receptor-related protein.  
ACCESSION X03362  
VERSION X03362.1 GI:56745  
KEYWORDS glycoprotein; kinase; neu oncogene; oncogene; transmembrane  
protein; tyrosine kinase.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
REFERENCE 1 (bases 1 to 3955)  
AUTHORS Bargmann, C.I., Hung, M.C. and Weinberg, R.A.  
TITLE The neu oncogene encodes an epidermal growth factor  
JOURNAL Nature 319 (6050), 226-230 (1986)  
FEATURES Location/Qualifiers



Klausner, R.D., Collins, F.S., Wagner, L., Shennen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,



BASE COUNT 1057 a 1346 c 1317 g 975 t

ORIGIN

Alignment Scores:

Pred. No.: 0.0653 Length: 4695  
Score: 84.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 52.17% Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x BC053078 (1-4695)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30

|||||

Db 3213 GTGGTCATCCAGACAGGAGCTTAGTCCCTCCAGCCCATGGACAGACCTTC 3266

RESULT 12

AY116182

LOCUS

DEFINITION Rattus norvegicus neu protooncoprotein mRNA, complete cds.

ACCESSION AY116182

VERSION AY116182.1 GI:22651764

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 4727)

AUTHORS Watson, P.A., Kim, K., Chen, K.-S. and Gould, M.N.

TITLE Androgen-Dependent Mammary Carcinogenesis in Rats Transgenic for

the New Proto-Oncogene

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 4727)

AUTHORS Chen, K.-S., Watson, P.A. and Gould, M.N.

TITLE Direct Submission

JOURNAL Submitted (29-MAY-2002) Oncology, University of Wisconsin, 1400

University Ave., Madison, WI 53706, USA

FEATURES Location/Qualifiers

source

1. 4727

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/strain="BDIX"

/db\_xref="taxon:10116"

106..3885

/note="p185 precursor; neu/erbB-2"

/codon\_start=1

/product="neu protooncoprotein"

/protein\_id="AA050093.1"

/db\_xref="gi:22651765"

/translation="MIMELAAACRWGFLALLPPGIAGTQVCTGTMKLRLPASPET

HLDMRLHYOGCOVQGNLETTYVPANASLSFLDIOEVQGYMLIAHNVKRVPLQLR

RIVRGTLQFEDKYLALAVLNDPNDVAASTPGRTPEGLRELQURSLTEILKGGVLIR

GNPOLCYDMWMDKVRKKNQLAPVDIDTNRSRACPPACPKDNHCWGESPEDCOI

LTGTCISGCARCKGRITPCDCEQCAAGCTGPKHSDCLACHFNHSGICELHCPALV

TYNTDTFESHNPBGRTYTFGASVYTCPTNYLSTFEGVCKTLCVCPNNQEVTAEDGTOR

CEKSKFCARVCYGLGHEGRLGARAITSDNVYDFGCKKIFGLSFLAFLPESFGDGPSSS

IAPLRPOLVFETLEETITGYLISAMPDLSRLDSVFQNLRIIEGRILHOGAVSLTIQ

GLGTHSLGSLRSLRELCSGLALHHRNHLCHFVHTVPMDLFRNPHQALLHSGNRPEEC

GLGLVNSLCAHGHCWGPPTQVCNCSHLRGQECVEECRWKGLPREYVSDKRCILP

CHPECPQNSSETCFGEADQCAACAHYKDSSCAVRCPSGVKPLDLYMPTIWKYPDE

GICQPCINFTSHCDVDERGAPAEORASPVFTIATVGVGLFLILVWVGLIKRR

GRURKTYMRLQLQETELVPLTPSGAMPNOAOMRILKETELRKVKVILGSGAFGTVK

GIWLPDGENKIPVAKIVLENTSPKANKELIDRAYVMAGVGSYPVSRILGICLTSTV

QLVTQLMPYCGLLDHRVHGRIGSQDLLNWCQVIAKGMSTYLEDVRLVRLHDAARNVL

TKSPNHVKITDFGLARLLDIDETEHADGGKVP IKWMALESILRRRTHSDVMSYGV

TVWELMTFGAKPDYGIIPAREIPDLLEKGERLPQPICTIDYIMVIMVCKWIMDSRCRP

FRELVSFSEARMARQFVIONEDLGPSPMDSTFYRSLLEDMDGDLVDAEYLVLP

QQGFSPDPPTGCTSTARHRRSSSTRSGGSELTLGLEPSEEGPPRPLAPSEAGSD

VFDGLAMGYTKGLQSLSPHDLSPLOKYSDEPTLPLPETDGYVAPLACSPQPEYVNO

SEVQPCPLTPPEGLPPVRPAGATLERPKTLSPGKNGVKDVAFAGAVENPEYLVPR

EGTASPPHPSPAFSPADNLYYWDQNSSEQGPPPSNFGTPTAENPEYLGIDVPV"

BASE COUNT 1004 a 1342 c 1355 g 1026 t

ORIGIN

Alignment Scores:

Pred. No.: 0.0658 Length: 4727  
Score: 84.00 Matches: 16  
Percent Similarity: 88.89% Conservative: 0  
Best Local Similarity: 88.89% Mismatches: 2  
Query Match: 52.17% Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ8\_SEQ2RES991\_ (1-30) x AY116182 (1-4727)

Qy 13 ValThrPheGlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30

|||||

Db 3076 GTGGTCATCCAGACAGGAGCTTGGGCCCATCCAGCCCATGGACAGTACCTTC 3129

RESULT 13

AL591390/c

LOCUS

DEFINITION Mouse DNA sequence from clone RP23-355L10 on chromosome 11,

complete sequence.

ACCESSION AL591390

VERSION AL591390.8 GI:17017767

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Phillimore, B.

REFERENCE 1 (bases 1 to 186134)

AUTHORS Direct Submission

TITLE

JOURNAL

Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquerry@sanger.ac.uk Clone requests: clonequest@sanger.ac.uk

On Nov 20, 2001 this sequence version replaced gi:16555512.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em, EMBL; Sw,

SWISSPROT; Tr, TREMBL; Wp, WORMPEP; information on the WORMPEP

database can be found at

http://www.sanger.ac.uk/Projects/C\_elegans/wormpep RP23-355L10 is

from the RPI-23 Mouse PAC Library

constructed by the group of Pieter de Jong.

For further details see http://www.chori.org/bacpac/home.htm

VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone

RP23-355L10 It may be shorter because we sequence overlapping

sections only once, except for a short overlap.

The true right end of clone RP23-355L10 is at 186134 in this

sequence. The true right end of clone RP23-438D7 is at 2000 in this

sequence.

FEATURES

source

1. 186134

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/chromosome="11"

/clone="RP23-355L10"

/clone\_lib="RPI-23"

17991..18044

misc\_feature





Job time : 465.208 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 234.794 Seconds  
(without alignments)  
3105.423 Million cell updates/sec

Title: JUNC\_SEQ3\_SEQ5

Perfect score: 159

Sequence: 1 DKGCPAQRASPLTSONEDIGPASPDLSTF 30

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO.spool\_P/HOLLERAN480/runat\_15102003\_131913\_20548/app\_query.fasta\_1.4685  
-DB=EST -QFMT=fastcap -SUFFIX=rrst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=HOLLERAN480 @CGN\_1\_1\_12645 @runat\_15102003\_131913\_20548 -ICPU=3  
-NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_etc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	85.5	53.8	618	28	AZ699313 RPCI-23-2
2	85	53.5	422	9	AW606421 RC3-HT037
3	85	53.5	439	10	BE157169 RC3-HT037
4	85	53.5	944	14	CD558559 AGENCOURT
5	81	50.9	458	10	BE157103 RC3-HT037
6	80	50.3	378	10	BF869854 IL3-EF011
7	80	50.3	411	2	BM069572 Homo sapi
8	80	50.3	433	10	BF762317 IL2-CS004
9	80	50.3	450	12	BG988031 MR2-HT116
10	80	50.3	466	12	BM720186 UI-E-E00-
11	80	50.3	546	12	BI194790 602948150
12	80	50.3	547	9	AA443351 zw85B05.i
13	80	50.3	561	12	BM720098 UI-E-E00-
14	80	50.3	607	12	BM083444 imagec-6
15	80	50.3	613	9	AW410534 fh06h06_x
16	80	50.3	666	13	BX113765 BX113765
17	80	50.3	667	12	BM048624 603623984
18	80	50.3	674	12	BM758528 K-EST0038
19	80	50.3	741	10	BE615590 601278973
20	80	50.3	881	9	AI906106 RC-BT105-
21	80	50.3	886	13	BQ882315 AGENCOURT
22	80	50.3	893	14	CA487631 AGENCOURT
23	80	50.3	905	13	BU541491 AGENCOURT
24	80	50.3	909	10	BG289457 602381495
25	80	50.3	922	14	CA453990 AGENCOURT
26	80	50.3	1009	13	BQ647343 AGENCOURT
27	79	49.7	278	10	BF923905 OVA-NT025
28	79	49.7	294	10	BF757857 CMA-CT057
29	79	49.7	337	10	BE091682 IL2-BT073
30	79	49.7	348	9	AI909847 QV-BT225-
31	79	49.7	378	14	CB266492 1005398 H
32	79	49.7	423	12	BM703962 UI-E-CK1-
33	79	49.7	464	2	BSM067221 Homo sapi
34	79	49.7	496	12	BM795191 K-EST0076
35	79	49.7	505	12	BM854421 K-EST0138
36	79	49.7	531	12	BM830170 K-EST0103
37	79	49.7	536	14	CB129403 K-EST0179
38	79	49.7	547	12	BM787824 K-EST0066
39	79	49.7	567	2	BSM07766 Homo sapi
40	79	49.7	574	12	BM829991 K-EST0103
41	79	49.7	630	2	BSM073298 Homo sapi
42	79	49.7	691	14	CB853376 UI-CF-FNO
43	79	49.7	767	10	BF240297 601905830
44	79	49.7	849	14	CA489799 AGENCOURT
45	79	49.7	852	14	CD516283 AGENCOURT

ALIGNMENTS

RESULT 1  
AZ699313  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE

AZ699313  
RPCI-23-240N16.TJ RPCI-23  
, genomic survey sequence.  
AZ699313  
A2699313.1 GI:12419623  
GSS.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 618)

618 bp  
DNA  
linear  
GSS 24-JAN-2001  
Mus musculus genomic clone RPCI-23-240N16

**AUTHORS** Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P., and Fraser, C. M.  
**TITLE** Mouse BAC End Sequences from Library RPCI-23  
**JOURNAL** Unpublished  
**COMMENT** Other\_GSSs: RPCI-23-240N16.TV  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
 Plate: 240 row: N column: 16  
 Seq primer: SP6  
 Class: BAC ends.

**FEATURES**

source

Location/Qualifiers

1..618  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-23-240N16"  
 /sex="Female"  
 /lab\_host="DH10B"  
 /clone\_lib="RPCI-23"  
 /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site.1: EcoRI; Site.2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methyase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 110 a 214 c 151 g 143 t

ORIGIN

## Alignment Scores:

Pred. No.: 0.347 Length: 618  
 Score: 85.50 Matches: 18  
 Percent Similarity: 72.41% Conservatives: 3  
 Best Local Similarity: 62.07% Mismatches: 5  
 Query Match: 53.77% Indels: 3  
 DB: 28 Gaps: 1

JUNC\_SEQ3\_SEQ5 (1-30) x AZ699313 (1-618)

QY 2 LysGlyCysProAlaGluArgAlaSerProLeuThrSerGlnAsnGluAspLeuGly 21  
 Db 165 GAGGGCTGTG-----CACCTCGTGTCTCCCTTCTCTCTAGACGAGACTTAGGC 215

QY 22 ProAlaSerProLeuAspSerThrPhe 30

Db 216 CCTCCAGCCCATGGACAGACCTTC 242

## RESULT 2

AW606421

LOCUS

RC3-HT0371-290100-013-e06 HT0371 Homo sapiens cDNA, linear EST 23-MAR-2000

DEFINITION

AW606421

ACCESSION

AW606421.1 GI:7311162

VERSION

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 422)

HCGP <http://www.ludwig.org.br/ORESTES>.

AUTHORS

TITLE

The FAPESP/LICR Human Cancer Genome Project

**JOURNAL**

COMMENT

Unpublished

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=RC3&t2=RC3-HT0371-290100-013-e06&t3=2000-01-29&t4=1>)

Seq primer: puc 18 forward

High quality sequence start: 8

High quality sequence stop: 422.

## FEATURES

source

Location/Qualifiers

1..422

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="HT0371"

/note="Organ: head/neck; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 77 a 126 c 135 g 84 t

ORIGIN

## Alignment Scores:

Pred. No.: 0.25 Length: 422  
 Score: 85.00 Matches: 23  
 Percent Similarity: 46.00% Conservatives: 0  
 Best Local Similarity: 46.00% Mismatches: 3  
 Query Match: 53.46% Indels: 24  
 DB: 9 Gaps: 2

JUNC\_SEQ3\_SEQ5 (1-30) x AW606421 (1-422)

QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15  
 Db 256 CCTGCAGAGGGTGGGAGGAGAGATGAGTCCAGTATGCCAGGCCCTCACGGAGGCTGC 315

QY 16 -----GlnAsnGluAspLeu 20  
 Db 316 ATGCTGGGCTGGGGAGGGGCCACCATCTGCTCTCTCTCCACAGAAATGAGGACTTG 375

QY 21 GlyProAlaSerProLeuAspSerThrPhe 30

Db 376 GGGCCAGCAGCTCCCTTGGACAGCACCTTC 405

## RESULT 3

BE157169/c

LOCUS

RC3-HT0371-250200-016-all HT0371 Homo sapiens cDNA, linear EST 21-JUN-2000

DEFINITION

BE157169

ACCESSION

BE157169.1 GI:8619890

VERSION

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 439)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

**TITLE** Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

**MEDLINE** 20202663

**PUBMED** 10737800

**COMMENT** Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-RC3-HT0371-250>)  
200-016-allt3-2000-02-256t4-1)  
Seq primer: puc 18 forward  
High quality sequence stop: 439.

# FEATURES

source

1. .439

Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="HT0371"

/notes="Organ: head,neck; Vector: puc18; Site.1: SmaI;

Site.2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

BASE COUNT 89 a 139 c 132 g 79 t

ORIGIN

# Alignment Scores:

Pred. No.: 0.262 Length: 439

Score: 85.00 Matches: 23

Percent Similarity: 46.00% Conservative: 0

Best Local Similarity: 46.00% Mismatches: 3

Query Match: 53.46% Indels: 24

DB: 10 Gaps: 2

JUNC\_SEQ3\_SEQ5 (1-30) x BE157169 (1-439)

QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15

Db 167 CCTGCAGAGGGTGGGAAGGAGAGATGAGTCCAGTATGCCAGGCCCTCACGGAAGGCTGC 108

QY 16 -----GlnAsnGluAspLeu 20

Db 107 ATGCTGGGCTGGGAGGGGCCACCACCTCTGCTCTCTCTCCACAGATGAGGACTTG 48

QY 21 GlyProAlaSerProLeuAspSerThrPhe 30

Db 47 GGCCCGACAGTCCCTTGGACAGCACCTTC 18

# RESULT 4

CD558559

LOCUS

DEFINITION

AGENCOURT\_14477527 NIH\_MGC\_181 Homo sapiens cDNA clone

IMAGE:30396701 5', mRNA sequence.

ACCESSION

CD558559

VERSION

CD558559.1 GI:31584627

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 944)

REFERENCE

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

# JOURNAL COMMENT

Unpublished

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

CDNA Library Prepared by: Dr. Michael Brownstein

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: NDAM478 row: n column: 06

High quality sequence start: 19

High quality sequence stop: 537.

FEATURES

source

1. .944

Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:30396701"

/tissue\_type="White Matter"

/dev\_stage="Unknown"

/lab\_host="NIH\_MGC\_181"

/clone\_lib="NIH\_MGC\_181"

/notes="Vector: pCMV-SPORT6, 1; Site.1: NotI; Site.2: EcoRV

(destroyed); Library is oligo-dT primed and directionally

cloned (EcoRV site is destroyed upon cloning). Average

insert size 1.42 kb. Library was constructed by

(Invitrogen). Note: this is a NIH\_MGC Library."

BASE COUNT 150 a 269 c 366 g 158 t

ORIGIN

1 others

# Alignment Scores:

Pred. No.: 0.682 Length: 944

Score: 85.00 Matches: 23

Percent Similarity: 46.00% Conservative: 0

Best Local Similarity: 46.00% Mismatches: 3

Query Match: 53.46% Indels: 24

DB: 14 Gaps: 2

JUNC\_SEQ3\_SEQ5 (1-30) x CD558559 (1-944)

QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15

Db 193 CCTGCAGAGGGTGGGAAGGAGAGATGAGTCCAGTATGCCAGGCCCTCACGGAAGGCTGC 252

QY 16 -----GlnAsnGluAspLeu 20

Db 253 ATGCTGGGCTGGGAGGGGCCACCACCTCTCTCTCTCTCCACAGATGAGGACTTG 312

QY 21 GlyProAlaSerProLeuAspSerThrPhe 30

Db 313 GGCCCGACAGTCCCTTGGACAGCACCTTC 342

# RESULT 5

BE157103

LOCUS

DEFINITION

RC3-HT0371-180200-015-h07 HT0371 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BE157103

VERSION

BE157103.1 GI:8619824

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 458)

REFERENCE

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsumura,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

20202663  
 10737800  
 PUBMED  
 COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IH3&t2=IL3-ET0114-251000-317-B07&t3=2000-10-25&t4=1>)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 177.

Location/Qualifiers  
 1..378  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="ET0114"  
 /note="Organ: lung\_tumor; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESPEC PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 76 a 119 c 110 g 73 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.944 Length: 378  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.31% Indels: 0  
 DB: 10 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x BF869854 (1-378)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
 |||||  
 Db 237 GACAAGGGTGTCCCGCCGAGCAGAGAGCCGCGCTTTGACGTC 193

RESULT 7  
 HSM069572  
 ID HSM069572 standard; RNA; EST; 411 BP.  
 XX  
 AC BX479682;  
 XX  
 SV BX479682.1  
 XX  
 DT 09-MAY-2003 (Rel. 75, Created)  
 DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)  
 XX  
 DE Homo sapiens mRNA; EST DKFZp686K212l3\_r1 (from clone DKFZp686K212l3)  
 XX  
 EST; expressed sequence tag.  
 XX  
 KW Homo sapiens (human)  
 XX OS  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia  
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 XX  
 [1]  
 RN 1-411  
 RP Bloembergen H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,  
 RA Han M., Wiemann S.;  
 RA Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.  
 RL MIMPS, Inqolstaedter Landstr.1, D-85764 Neuherberg, GERMANY  
 RL



```

XX CC This is the 5' sequence of the clone insert
CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
CC sequenced by GBF (National Research Centre for Biotechnology
CC Ltd., Braunschweig/Germany) within the cDNA sequencing
CC consortium of the German Genome Project.
CC No s1 sequence available.
CC This clone (DKFZp686K21213) is available at the RZPD in Berlin.
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,
CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX FH
XX Key Location/Qualifiers
FH source
FT 1. 411
FT /db_xref="taxon:9606"
FT /mol_type="mRNA"
FT /organism="Homo sapiens"
FT /clone="DKFZp686K21213"
FT /clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
FT DH10B; sites SfiIA + SfiIB"
FT /dev_stage="adult"
FT /tissue_type="CDNA-collection"
XX SQ Sequence 411 BP; 70 A; 126 C; 125 G; 90 T; 0 other;

Alignment Scores:
Pred. No.: 1.05 Length: 411
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x HSM069572 (1-411)

QY 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
DB 292 GACAAGGCTGCCCGCCGAGCAGAGACCCCTCTGACGTCC 336

RESULT 8
BF762317/c 433 bp mRNA linear EST 12-JAN-2001
LOCUS IL2-CS0049-181000-202-A06 CS0049 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF762317
ACCESSION BF762317
VERSION BF762317.1 GI:12110217
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 433)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR2&t2=MR2-HT1160-
110101-005-a06&t3=2001-01-11&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 39
High quality sequence stop: 325.

```

```

(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=IL2&t2=IL2-CS0049-
181000-202-A06&t3=2000-10-18&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 387.
Location/Qualifiers
1. 433
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CS0049"
/Note="Organ: colon_est; Vector: puc18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 86 a 149 c 108 g 90 t
ORIGIN

Alignment Scores:
Pred. No.: 1.12 Length: 433
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x BF762317 (1-433)

QY 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
DB 367 GACAAGGCTGCCCGCCGAGCAGAGACCCCTCTGACGTCC 323

RESULT 9
BG988031 450 bp mRNA linear EST 13-JUN-2001
LOCUS MR2-HT1160-110101-005-a06 HT1160 Homo sapiens cDNA, mRNA sequence.
DEFINITION BG988031
ACCESSION BG988031
VERSION BG988031.1 GI:14392101
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 450)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR2&t2=MR2-HT1160-
110101-005-a06&t3=2001-01-11&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 39
High quality sequence stop: 325.

```

```

FEATURES
source
  Location/Qualifiers
    1..450
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /dev_stage="Adult"
      /clone_lib="HT1160"
      /note="Organ: head,neck; Vector: puc18; Site_1: SmaI;
      Site_2: SmaI; A mini-library was made by cloning products
      derived from ORESTES PCR (U.S. Letters Patent application
      No. 196,716 - Ludwig Institute for Cancer Research)
      profiles into the pUC 18 vector. Reverse transcription of
      tissue mRNA and cDNA amplification were performed under
      low stringency conditions."
BASE COUNT      84 a 127 c 150 g 89 t
ORIGIN
Alignment Scores:
Pred. No.:      1.17      Length:      450
Score:          80.00      Matches:      15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      50.31%      Indels:      0
DB:              12      Gaps:
JUNC_SEQ3_SEQ5 (1-30) x BG988031 (1-450)
QY      1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db      259 GACAAGGGCTGCCCTCGCAGCAGAGAGCCAGCCCTCTGACGTC 303

RESULT 10
BM720186
LOCUS      BM720186      466 bp      mRNA      linear      EST 01-MAR-2002
DEFINITION UI-E-E00-ahz-c-08-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone
ACCESSION      BM720186
VERSION      BM720186.1 GI:19039265
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1 (bases 1 to 466)
AUTHORS      Bonaudo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      97044477
PubMed      8889548
COMMENT      Contact: Soares, MB
      Coordinated Laboratory for Computational Genomics
      University of Iowa
      375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
      Tel: 319 335 8250
      Fax: 319 335 9565
      Email: bento-soares@uiowa.edu
      Tissue Procurement: Dr. Gregg Hageman
      CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
      CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
      DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
      Clone Distribution: Researchers may obtain clones from Research
      Genetics (www.resgen.com).
      Seq primer: M13 Reverse.
      Location/Qualifiers
        1..466
          /organism="Homo sapiens"
          /mol_type="mRNA"
          /db_xref="taxon:9606"
          /clone="UI-E-E00-ahz-c-08-0-UI"
          /tissue_type="fetal eye"
          /dev_stage="fetal"
          /lab_host="DH10B (Life Technologies) (T1 phage resistant)"

FEATURES
source
  Location/Qualifiers
    1..546
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:5091396"
      /tissue_type="epithelioid carcinoma cell line"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH_MGC_42"
      /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
      Site_2: EcoRI; cDNA made by oligo-dT priming.
      Directionally cloned into EcoRI/XhoI sites using the

```

```

/clone_lib="UI-E-E00"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-E00 is a cDNA library containing the following
tissue(s): fetal eye. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CGGTATACC. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."
BASE COUNT      105 a 114 c 151 g 93 t 3 others
ORIGIN
Alignment Scores:
Pred. No.:      1.23      Length:      466
Score:          80.00      Matches:      15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      50.31%      Indels:      0
DB:              12      Gaps:
JUNC_SEQ3_SEQ5 (1-30) x BM720186 (1-466)
QY      1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db      18 GACAAGGGCTGCCCTCGCAGCAGAGCCAGCCCTCTGACGTC 62

RESULT 11
BI194790
LOCUS      BI194790      546 bp      mRNA      linear      EST 10-JUL-2001
DEFINITION 602948150F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:5091396 5',
mRNA sequence.
ACCESSION      BI194790
VERSION      BI194790.1 GI:14649810
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1 (bases 1 to 546)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
      Email: cgapbs-r@mail.nih.gov
      Tissue Procurement: ATCC
      CDNA Library Preparation: Ling Hong/Rubin Laboratory
      CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)
      DNA Sequencing by: Incyte Genomics, Inc.
      Clone Distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/LEML at:
      http://image.llnl.gov
      Plate: LCM1855 row: c column: 13
      High quality sequence start: 10
      High quality sequence stop: 495.
FEATURES
source
  Location/Qualifiers
    1..546
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:5091396"
      /tissue_type="epithelioid carcinoma cell line"
      /lab_host="DH10B (phage-resistant)"
      /clone_lib="NIH_MGC_42"
      /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
      Site_2: EcoRI; cDNA made by oligo-dT priming.
      Directionally cloned into EcoRI/XhoI sites using the

```

following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC library. |"

BASE COUNT 110 a 147 c 183 g 106 t

ORIGIN

Alignment Scores:  
Pred. No.: 1.49 Length: 546  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x B1194790 (1-546)

Qy 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 223 GACAAGGGCTGCCCCCGGAGCAGAGCCGCCCTCTGACGTC 267

RESULT 12

AA443351

LOCUS

DEFINITION

zw85b05.r1 Soares total fetus Nb2HF8.9w Homo sapiens linear EST 03-JUN-1997  
IMAGE:783729 5' similar to gb:U11730 ERBB-2 RECEPTOR  
PROTEIN-TYROSINE KINASE PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION

AA443351

VERSION

AA443351.1 GI:2156026

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 547)

ILLIER, L., ALLEN, M., BOWLES, L., DUBUCQUE, T., GEISEL, G., JOST, S.,

KUCABA, T., LACY, M., LE, N., LENNON, G., MARA, M., MARTIN, J., MOORE, B.,

SCHILLENBERG, K., STEPTOE, M., TAN, F., THEISING, B., WHITE, Y., WYLIE

, T., WATERSTON, R., AND WILSON, R.

WASHU-MERCK EST Project 1997

Unpublished

Contact: Wilton RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 360.

Location/Qualifiers

1..547

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:5981347"

/db\_xref="taxon:9606"

/clone="IMAGE:783729"

/dev\_stage="8-9 weeks"

/lab\_host="DH10B"

/clone\_lib="Soares total fetus Nb2HF8.9w"

/note="Vector: pT7T3-Pac (Pharmacia) with a modified

polylinker; Site: Not I; Site: Eco RI; 1st strand cDNA

was prepared from mRNA obtained from pooled 8-9 week

(total) fetus material with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGAGCGCGCTTAATTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT7T3 vector. Library

went through one round of normalization, and was

constructed by Bento Soares and M. Fatima Bonaldo. "

FEATURES

source

BASE COUNT 115 a 151 c 172 g 108 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 1.5 Length: 547  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 9 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x AA443351 (1-547)

Qy 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 174 GACAAGGGCTGCCCCCGGAGCAGAGCCGCCCTCTGACGTC 218

RESULT 13

BM720098

LOCUS

DEFINITION

UI-E-E00-ahz-c-05-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone

UI-E-E00-ahz-c-05-0-UI 5', mRNA sequence.

ACCESSION

BM720098

VERSION

BM720098.1 GI:19039119

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 561)

BONALDO, M.F., LENNON, G. AND SOARES, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

8889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

Location/Qualifiers

1..561

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="UI-E-E00-ahz-c-05-0-UI"

/tissue\_type="fetal eye"

/dev\_stage="fetal"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone\_lib="UI-E-E00"

/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site: 1: EcoR I; Site: 2: Not I;

UI-E-E00 is a cDNA library containing the following

tissue(s): fetal eye. The library was constructed

according to Bonaldo, Lennon and Soares, Genome Research,

6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dT primer containing a Not I site. Double

stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pT7T3-Pac

vector. The oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The

sequence tag for this library is GCGGTATACC. This library

was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

```

BASE COUNT      97 a 171 c 179 g 113 t 1 others
ORIGIN

Alignment Scores:
Pred. No.:      1.54      Length:      561
Score:          80.00     Matches:      15
Percent Similarity: 100.00%  Conservaive: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    50.31%   Indels:      0
DB:            12      Gaps:      0

JUNC_SEQ3_SEQ5 (1-30) x BM720098 (1-561)

QY      1 AsPLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db      400 GACAAGGCGTGCCTCCGCGAGAGAGAGCCAGCCCTCTGAGTCC 444

RESULT 14
BM083444
LOCUS      BM083444      607 bp      mRNA      linear      EST 16-NOV-2001
DEFINITION imagec_6_2000/sjpa59bdf42.x2 NIH_MGC_39 Homo sapiens cDNA clone
ACCESSION  BM083444
VERSION     BM083444.1 GI:16951075
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           Kale, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and
           Prange, C.K.
TITLE       The I.M.A.G.E. Consortium quality control effort: clone
JOURNAL    Resequencing for verification
COMMENT    Other_ESTs: BE615590
           Contact: Prange CK
           The I.M.A.G.E. Consortium
           Lawrence Livermore National Laboratory
           Livermore, CA, USA
           Email: help@image.llnl.gov
           This read has been verified (found to hit its original self in the
           correct orientation), as part of the I.M.A.G.E. Consortium quality
           control effort. High quality sequence is defined as having 100 or
           more base pairs with a phred quality value of 20 or greater, where
           a sliding window of 4 base pairs with a phred quality value of 15
           or greater marks the beginning and end of the sequence. For
           information on obtaining this clone, please contact
           info@image.llnl.gov.
           Plate: L1CM267 row: b column: 19
           Seq primer: -21ml3
           High quality sequence stop: 607.

FEATURES             source
     1..607
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:3610674"
     /tissue_type="adenocarcinoma"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH_MGC_39"
     /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
     Site_2: EcoRI; cDNA made by oligo-dT priming.
     Directionally cloned into EcoRI/XhoI sites using the
     following 5' adaptor: GGCACGAG(G). Library constructed
     by Ling Hong in the laboratory of Gerald M. Rubin
     (University of California, Berkeley) using ZAP-cDNA
     synthesis kit (Stratagene) and Superscript II RT (Life
     Technologies)."

BASE COUNT      118 a 171 c 195 g 121 t 2 others
ORIGIN

```

## Alignment Scores:

Pred. No.: 1.7 Length: 607  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00% Conservaive: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.31% Indels: 0  
 DB: 12 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x BM083444 (1-607)

QY 1 AsPLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
 |||||  
 Db 292 GACAAGGCGTGCCTCCGCGAGAGAGCCAGCCCTCTGAGTCC 335

## RESULT 15

AW410534 613 bp mRNA linear EST 29-JUN-2000  
 LOCUS fh06h06.x1 NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:2961635 5',  
 DEFINITION mRNA sequence.  
 ACCESSION AW410534  
 VERSION AW410534.1 GI:6936075  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 REFERENCE 1 (bases 1 to 613)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: coapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bbrp/image/image.html  
 Plate: L1CM36 row: 0 column: 12  
 Seq primer: -21M13 forward primer (ABI).  
 Location/Qualifiers  
 1..613  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2961635"  
 /tissue\_type="rhabdomyosarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_17"  
 /note="Organ: muscle; Vector: pOTB7; Site\_1: EcoRI;  
 Site\_2: XhoI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

## FEATURES

## source

1..613  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2961635"  
 /tissue\_type="rhabdomyosarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_17"  
 /note="Organ: muscle; Vector: pOTB7; Site\_1: EcoRI;  
 Site\_2: XhoI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

## BASE COUNT

## ORIGIN

## Alignment Scores:

Pred. No.: 1.73 Length: 613  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00% Conservaive: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.31% Indels: 0  
 DB: 9 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x AW410534 (1-613)

Qy 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||  
Db 557 GACAAGGCTGCCCGCGGAGCAGAGAGCCAGCCCTCTGACGTCC 601

Search completed: October 16, 2003, 17:04:03  
Job time : 235.794 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 16, 2003, 11:08:07 ; Search time 31.0111 Seconds  
(without alignments)  
2540.503 Million cell updates/sec

Title: JUNC\_SEQ3\_SEQ5

Perfect score: 159

Sequence: 1 DKGCPARQASPLTSQNEGLGASPLDSTF 30

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+lp2n.model -DEV=xlp  
-Q/cgn2\_1/USFPO\_spool\_p/HOLLERAN480/runat\_15102003\_131915\_20662/app\_query.fasta\_1.4685  
-DB=Published Applications NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blotsum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=HOLLERAN480 @CGN\_1\_1\_1\_397 @runat\_15102003\_131915\_20662  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	80	50.3	201 13	US-10-109-213-3 Sequence 3, Appli

2	80	50.3	3765	12	US-10-207-498-5	Sequence 5, Appli
3	80	50.3	3768	9	US-09-811-123-8	Sequence 8, Appli
4	80	50.3	3768	9	US-09-811-115-2	Sequence 2, Appli
5	80	50.3	3768	10	US-09-854-356-9	Sequence 9, Appli
6	80	50.3	3768	10	US-09-930-125-1	Sequence 1, Appli
7	80	50.3	3768	12	US-10-313-644-1	Sequence 1, Appli
8	80	50.3	4473	11	US-09-441-411-5	Sequence 5, Appli
9	80	50.3	4473	12	US-10-101-510-81	Sequence 81, Appli
10	80	50.3	4473	14	US-10-146-473-32	Sequence 32, Appli
11	80	50.3	4473	14	US-10-207-655-44	Sequence 44, Appli
12	80	50.3	4530	10	US-09-877-177-11	Sequence 11, Appli
13	80	50.3	4530	12	US-10-007-926A-119	Sequence 70, Appli
14	80	50.3	4530	12	US-10-101-510-124	Sequence 119, App
15	80	50.3	4530	12	US-10-338-730-1	Sequence 124, App
16	80	50.3	4530	14	US-10-177-293-125	Sequence 1, Appli
17	80	50.3	4543	10	US-09-769-508-1	Sequence 125, App
18	80	50.3	4606	12	US-09-971-392-70	Sequence 1, Appli
19	80	50.3	4642	14	US-10-198-846-10896	Sequence 10896, A
20	80	50.3	9274	9	US-09-811-123-7	Sequence 7, Appli
21	80	50.3	9274	9	US-09-811-115-1	Sequence 1, Appli
22	79	49.7	1713	12	US-10-102-806-165	Sequence 165, App
23	79	49.7	1713	12	US-10-378-393-14	Sequence 14, Appli
24	79	49.7	1755	10	US-09-930-125-6	Sequence 6, Appli
25	79	49.7	1767	10	US-09-930-125-4	Sequence 4, Appli
26	79	49.7	1773	10	US-09-930-125-7	Sequence 7, Appli
27	79	49.7	1806	10	US-09-930-125-5	Sequence 5, Appli
28	79	49.7	2411	12	US-10-378-393-10	Sequence 10, Appli
29	74	46.5	3771	10	US-09-854-356-11	Sequence 11, Appli
30	74	46.5	3955	10	US-09-870-759-117	Sequence 117, App
31	74	46.5	3955	10	US-09-854-356-10	Sequence 10, Appli
32	74	46.5	3955	12	US-09-751-708A-117	Sequence 117, App
33	62	39.0	14427	14	US-10-156-761-1540	Sequence 1540, Ap
34	62	39.0	23432	9	US-09-764-969-1332	Sequence 1332, Ap
35	62	39.0	23432	14	US-10-091-504-1332	Sequence 1332, Ap
36	62	39.0	9025608	14	US-10-156-761-1	Sequence 1, Appli
37	59	37.1	111282	13	US-10-094-989-3	Sequence 3, Appli
38	58.5	36.8	633	11	US-09-788-188-12	Sequence 12, Appli
39	58.5	36.8	634	8	US-08-450-842-1	Sequence 1, Appli
40	58.5	36.8	1404	14	US-10-155-785-13	Sequence 13, Appli
41	57.5	36.2	363	11	US-09-918-995-28706	Sequence 28706, A
42	57	35.8	297	10	US-09-783-590-11498	Sequence 11498, A
43	57	35.8	1938	12	US-10-422-264-25	Sequence 25, Appli
44	57	35.8	2316	12	US-10-422-264-27	Sequence 27, Appli
45	57	35.8	2604	12	US-10-422-264-23	Sequence 23, Appli

ALIGNMENTS

RESULT 1

US-10-109-213-3  
; Sequence 3, Application US/10109213  
; Publication No. US20020186670A1  
; GENERAL INFORMATION:  
; APPLICANT: Ecker, David J.  
; TITLE OF INVENTION: Identification Of Disease Predictive Nucleic Acids  
; FILE REFERENCE: IBIS0009  
; CURRENT APPLICATION NUMBER: US/10/109,213  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: US/09/200,355  
; PRIOR FILING DATE: 1998-11-25  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 201  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-10-109-213-3

Alignment Scores:	0.000724	Length:	201
Pred. No.:	80.00	Matches:	15
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%		

Query Match: 50.31% Indels: 0  
DB: 13 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x US-10-109-213-3 (1-201)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 116 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCUCUGACGTC 160

## RESULT 2

US-10-207-498-5  
; Sequence 5, Application US/10207498  
; Publication No. US20030143568A1  
; GENERAL INFORMATION:  
; APPLICANT: Elizabeth Singer  
; APPLICANT: Ralf Landgraf  
; APPLICANT: Dennis J. Slamon  
; APPLICANT: David Eisenberg  
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING  
; FILE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HEREGULIN AND HER3  
; FILE REFERENCE: 30448.103-US-U1  
; CURRENT APPLICATION NUMBER: US/10/207,498  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR FILING DATE: 2001-07-27  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 3765  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(3765)  
US-10-207-498-5

Alignment Scores:  
Pred. No.: 0.0177 Length: 3765  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x US-10-207-498-5 (1-3765)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTC 1959

## RESULT 3

US-09-811-123-8  
; Sequence 8, Application US/09811123  
; Patent No. US20020001587A1  
; GENERAL INFORMATION:  
; APPLICANT: Sharon Erickson  
; APPLICANT: Ralph Schwall  
; APPLICANT: Mark Sliwowski  
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ErbB  
; FILE OF INVENTION: ANTIBODY-MAYTANSINOID CONJUGATES  
; FILE REFERENCE: GENENT.073A2  
; CURRENT APPLICATION NUMBER: US/09/811,123  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 60/238,327  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: 09/602,530  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 3768  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-811-123-8

Alignment Scores:  
Pred. No.: 0.0177 Length: 3768  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 9 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x US-09-811-123-8 (1-3768)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTC 1959

## RESULT 4

US-09-811-115-2  
; Sequence 2, Application US/09811115  
; Patent No. US20020035736A1  
; GENERAL INFORMATION:  
; APPLICANT: Erickson, Sharon  
; APPLICANT: Schwall, Ralph  
; APPLICANT: King, Kathleen  
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL  
; FILE REFERENCE: GENENT.034A  
; CURRENT APPLICATION NUMBER: US/09/811,115  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR FILING DATE: 2000-03-16  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 3768  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-811-115-2

Alignment Scores:  
Pred. No.: 0.0177 Length: 3768  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 9 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x US-09-811-115-2 (1-3768)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTC 1959

## RESULT 5

US-09-854-356-9  
; Sequence 9, Application US/09854356  
; Patent No. US20020177567A1  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Gheysen, Dirk  
; APPLICANT: Corixa Corporation  
; APPLICANT: SmithKline Beecham Biologicals S. A.  
; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
; FILE REFERENCE: 014058-009810PC  
; CURRENT APPLICATION NUMBER: US/09/854,356  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 09/493,480  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/117,976  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 3768



```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2076)..(3765)
; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9
```

```
Alignment Scores:
Pred. No.: 0.0177 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 10 Gaps: 0
```

JUNC\_SEQ3\_SEQ5 (1-30) x US-09-854-356-9 (1-3768)

```
Qy 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGCTGCCCGCGGACGAGAGCCAGCCCTCTGACGTCC 1959
```

## RESULT 6

```
US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930.125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3765)
US-09-930-125-1
```

```
Alignment Scores:
Pred. No.: 0.0177 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 10 Gaps: 0
```

JUNC\_SEQ3\_SEQ5 (1-30) x US-09-930-125-1 (1-3768)

```
Qy 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGCTGCCCGCGGACGAGAGCCAGCCCTCTGACGTCC 1959
```

## RESULT 7

```
US-10-313-644-1
; Sequence 1, Application US/10313644
; Publication No. US20030157119A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: Hand-Zimmerman, Susan
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
; OTHER INFORMATION: AND VIRUS-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.483C3
; CURRENT APPLICATION NUMBER: US/10/313.644
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3765)
US-10-313-644-1
```

```
Alignment Scores:
Pred. No.: 0.0177 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 12 Gaps: 0
```

JUNC\_SEQ3\_SEQ5 (1-30) x US-10-313-644-1 (1-3768)

```
Qy 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGCTGCCCGCGGACGAGAGCCAGCCCTCTGACGTCC 1959
```

## RESULT 8

```
US-09-441-411-5
; Sequence 5, Application US/09441411
; Publication No. US20030008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441.411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-5
```

```
Alignment Scores:
Pred. No.: 0.0214 Length: 4473
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 11 Gaps: 0
```

JUNC\_SEQ3\_SEQ5 (1-30) x US-09-441-411-5 (1-4473)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 2089 GACAAGGGCTGCCCGCGAGAGAGAGCCAGCCCTCTGACGTCC 2133

## RESULT 9

US-10-101-510-81  
; Sequence 81, Application US/10101510  
; Publication No. US20030148295A1  
; GENERAL INFORMATION:  
; APPLICANT: WAN, JACKSON  
; APPLICANT: WANG, YIXIN  
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE  
; FILE REFERENCE: 15117.0012  
; CURRENT APPLICATION NUMBER: US/10/101,510  
; CURRENT FILING DATE: 2002-03-20  
; PRIOR APPLICATION NUMBER: 60/276,947  
; PRIOR FILING DATE: 2001-03-20  
; NUMBER OF SEQ ID NOS: 805  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 81  
; LENGTH: 4473  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-101-510-81

Alignment Scores:  
Pred. No.: 0.0214 Length: 4473  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x US-10-101-510-81 (1-4473)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 2089 GACAAGGGCTGCCCGCGAGAGAGCCAGCCCTCTGACGTCC 2133

## RESULT 10

US-10-146-473-32  
; Sequence 32, Application US/10146473  
; Publication No. US20030108888A1  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Matthew  
; APPLICANT: Gout, Ivan  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Gure, Ali  
; APPLICANT: Chen, Yao-Tseng  
; APPLICANT: Old, Lloyd  
; TITLE OF INVENTION: Breast Cancer Antigens  
; FILE REFERENCE: L00461/70130(JRV)  
; CURRENT APPLICATION NUMBER: US/10/146,473  
; CURRENT FILING DATE: 2002-05-15  
; PRIOR APPLICATION NUMBER: US 60/291,150  
; PRIOR FILING DATE: 2001-05-15  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 32  
; LENGTH: 4473  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-146-473-32

Alignment Scores:  
Pred. No.: 0.0214 Length: 4473  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 14 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x US-10-146-473-32 (1-4473)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 2089 GACAAGGGCTGCCCGCGAGAGAGCCAGCCCTCTGACGTCC 2133

## RESULT 11

US-10-207-655-44  
; Sequence 44, Application US/10207655  
; Publication No. US20030118592A1  
; GENERAL INFORMATION:  
; APPLICANT: Ledbetter, Jeffrey A.  
; APPLICANT: Hayden-Ledbetter, Martha S.  
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
; FILE REFERENCE: 390069.401C1  
; CURRENT APPLICATION NUMBER: US/10/207,655  
; CURRENT FILING DATE: 2002-07-25  
; NUMBER OF SEQ ID NOS: 426  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 44  
; LENGTH: 4473  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-207-655-44

Alignment Scores:  
Pred. No.: 0.0214 Length: 4473  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 14 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x US-10-207-655-44 (1-4473)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 2089 GACAAGGGCTGCCCGCGAGAGAGCCAGCCCTCTGACGTCC 2133

## RESULT 12

US-09-877-177-11  
; Sequence 11, Application US/09877177  
; Publication No. US20020192652A1  
; GENERAL INFORMATION:  
; APPLICANT: Peter V. Danenberg et al.  
; TITLE OF INVENTION: Method of determining Epidermal Growth  
; FACTOR RECEPTOR AND HER2-NEW GENE EXPRESSION  
; FILE REFERENCE: 11220/120  
; CURRENT APPLICATION NUMBER: US/09/877,177  
; CURRENT FILING DATE: 2001-06-11  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 4530  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-877-177-11

Alignment Scores:  
Pred. No.: 0.0217 Length: 4530  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x US-09-877-177-11 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 2065 GACAAGGGCTGCCCGCGAGAGAGCCAGCCCTCTGACGTCC 2109

## RESULT 13

US-10-007-926A-119

```

; Sequence 119, Application US/10007926A
; Publication No. US20030143539A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007,926A
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 119
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: v-erb-b2 avian erythroblastic leukemia viral
; OTHER INFORMATION: oncogene homolog 2 (neuro/glioblastoma derived
; OTHER INFORMATION: oncogene homolog) (ERBB2) gene.
US-10-007-926A-119

Alignment Scores:
Pred. No.:      0.0217          Length:      4530
Score:           80.00          Matches:       15
Percent Similarity: 100.00%    Conservative:   0
Best Local Similarity: 100.00% Mismatches:     0
Query Match:      50.31%      Indels:        0
DB:               12          Gaps:         0

JUNC_SEQ3_SEQ5 (1-30) x US-10-007-926A-119 (1-4530)

Qy      1 AsPLysGLYCysProAlaGlucLnArgAlaSerProLeuThrSer 15
Db      2065 GACAAGGGTGTCCCCCGCCGACGAGACGCCAGCCCCTCTGACGTC 2109
|||||
RESULT 14
US-10-101-510-124
; Sequence 124, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 124
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-124

Alignment Scores:
Pred. No.:      0.0217          Length:      4530
Score:           80.00          Matches:       15
Percent Similarity: 100.00%    Conservative:   0
Best Local Similarity: 100.00% Mismatches:     0
Query Match:      50.31%      Indels:        0
DB:               12          Gaps:         0

JUNC_SEQ3_SEQ5 (1-30) x US-10-101-510-124 (1-4530)

```

**THIS PAGE BLANK (USPTO)**



```
;
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-09-146-283-3

Alignment Scores:
Pred. No.: 0.0053 Length: 2385
Score: 84.00 Matches: 21
Percent Similarity: 70.00% Conservative: 0
Best Local Similarity: 70.00% Mismatches: 5
Query Match: 52.83% Indels: 4
DB: 2 Gaps: 2

JUNC_SEQ3_SEQ5 (1-30) x US-09-146-283-3 (1-2385)
QY 1 ASPLYSGLYCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1925 GACAAGGCGTGCCTCCCGCAGCAGAGCCCTCTGACGTCCTCGAG----- 1975
QY 21 GlyProAla---SerProLeuAspSerThr 29
Db 1976 GCACCGCGCGCTCGCCAGCCCGCCAGCACA 2005

RESULT 2
US-08-579-823A-3
; Sequence 3, Application US/08579823A
; Patent No. 6080409
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Composition and Method
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,823A
; FILING DATE: 03-DEC-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-344-195-3

Alignment Scores:
Pred. No.: 0.0053 Length: 2385
Score: 84.00 Matches: 21
Percent Similarity: 70.00% Conservative: 0
Best Local Similarity: 70.00% Mismatches: 5
Query Match: 52.83% Indels: 4
DB: 2 Gaps: 2

JUNC_SEQ3_SEQ5 (1-30) x US-08-579-823A-3 (1-2385)
QY 1 ASPLYSGLYCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1925 GACAAGGCGTGCCTCCCGCAGCAGAGCCCTCTGACGTCCTCGAG----- 1975
QY 21 GlyProAla---SerProLeuAspSerThr 29
Db 1976 GCACCGCGCGCTCGCCAGCCCGCCAGCACA 2005

RESULT 3
US-09-344-195-3
; Sequence 3, Application US/09344195
; Patent No. 6210662
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,195
; FILING DATE: 24-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-344-195-3

Alignment Scores:
```

Pred. No.: 0.0053 Length: 2385  
Score: 84.00 Matches: 21  
Percent Similarity: 70.00% Conservative: 0  
Best Local Similarity: 70.00% Mismatches: 5  
Query Match: 52.83% Indels: 4  
DB: 3 Gaps: 2

JUNC\_SEQ3\_SEQ5 (1-30) x US-09-344-195-3 (1-2385)

Qy 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20  
Db 1925 GACAAAGGCTGCCCGCGAGAGAGCCAGCCCTCTGAGCTCCCTCGAG----- 1975  
Qy 21 GlyProAla---SerProLeuAspSerThr 29  
Db 1976 GCACCGCGCTCGCCGAGCCAGCCAGCACA 2005

## RESULT 4

US-08-776-251-3  
; Sequence 3, Application US/08776251  
; Patent No. 6025340  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Caroline J  
; APPLICANT: Marais, Richard  
; TITLE OF INVENTION: Surface expression of enzyme in gene directed prodrug therapy  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye  
; STREET: 1100 No. 6025340th Glebe Road, 8th Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: USA  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,251  
; FILING DATE: 31-JAN-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB95/01782  
; FILING DATE: 27-JUL-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9415167.7  
; FILING DATE: 27-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arthur R. Crawford  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 620-20  
; INFORMATION FOR SEQ ID NO: 3:  
; LENGTH: 153 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
US-08-776-251-3

Alignment Scores:  
Pred. No.: 0.000654 Length: 153  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 3 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x US-08-776-251-3 (1-153)

Qy 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15  
Db 10 GACAAAGGCTGCCCGCGAGAGAGCCAGCCCTCTGAGCTCC 54

RESULT 5  
US-09-200-355-3  
; Sequence 3, Application US/09200355  
; Patent No. 6451524  
; GENERAL INFORMATION:  
; APPLICANT: Ecker, David J.  
; TITLE OF INVENTION: Identification of Disease Predictive Nucleic Acids  
; FILE REFERENCE: IBIS0009  
; CURRENT APPLICATION NUMBER: US/09/200,355  
; CURRENT FILING DATE: 1998-11-25  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 201  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-200-355-3

Alignment Scores:  
Pred. No.: 0.000922 Length: 201  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 4 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x US-09-200-355-3 (1-201)

Qy 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15  
Db 116 GACAAGGCGGCCCGCGAGAGAGCCAGCCCGCUCGAGCGGCC 160

## RESULT 6

US-08-776-251-10  
; Sequence 10, Application US/08776251  
; Patent No. 6025340  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Caroline J  
; APPLICANT: Marais, Richard  
; TITLE OF INVENTION: Surface expression of enzyme in gene directed prodrug therapy  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye  
; STREET: 1100 No. 6025340th Glebe Road, 8th Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: USA  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,251  
; FILING DATE: 31-JAN-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB95/01782  
; FILING DATE: 27-JUL-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9415167.7  
; FILING DATE: 27-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arthur R. Crawford  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 620-20  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 816 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
US-08-776-251-10

```

Alignment Scores:
Pred. No.: 0.00536 Length: 816
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 3 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-08-776-251-10 (1-816)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 93 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCTCTGACGTCC 137

RESULT 7
US-08-625-101-1
; Sequence 1, Application US/08625101
; Patent No. 5869445
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,101
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3765
US-08-625-101-1

Alignment Scores:
Pred. No.: 0.0366 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-08-625-101-1 (1-3768)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||

```

```

Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCTCTGACGTCC 1959

RESULT 8
US-08-356-786-1
; Sequence 1, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3768
; OTHER INFORMATION: /note= "product = "cerB-b2""
US-08-356-786-1

Alignment Scores:
Pred. No.: 0.0366 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-08-356-786-1 (1-3768)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCTCTGACGTCC 1959

RESULT 9
US-09-048-804-1
; Sequence 1, Application US/09048804
; Patent No. 5968748
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters

```



Pred. No.: 0.0454  
Length: 4473

GENERAL INFORMATION:

```
;
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-645-865-9

Alignment Scores:
Pred. No.: 0.0462 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-08-645-865-9 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2065 GACAAGGGCTGCCCGCCGAGCAGAGCCGCTCTGACGTCC 2109

RESULT 13
US-09-167-322-4
; Sequence 4, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
```

```
;
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-167-322-4

Alignment Scores:
Pred. No.: 0.0462 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-09-167-322-4 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2065 GACAAGGGCTGCCCGCCGAGCAGAGCCGCTCTGACGTCC 2109

RESULT 14
US-09-527-487-1
; Sequence 1, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(3915)
US-09-527-487-1

Alignment Scores:
Pred. No.: 0.0462 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x US-09-527-487-1 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 2065 GACAAGGGCTGCCCGCCGAGCAGAGCCGCTCTGACGTCC 2109

RESULT 15
US-09-877-177A-11
; Sequence 11, Application US/09877177A
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:06:28 : Search time 28.4941 Seconds  
(without alignments)  
2842.104 Million cell updates/sec

Title: JUNC\_SEQ3\_SEQ5  
Perfect score: 159  
Sequence: 1 DKGCPAEQRAPLTSQNDLGPASPLDSTF 30

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool.p/HOLLERAN480/runat\_15102003\_131912\_20526/app\_query.fasta\_1.4685  
-DB=N\_Geneseq\_19Jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptlo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 -CGN\_1\_1\_1758 @runat\_15102003\_131912\_20526 -NCFU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25:	/SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	144	90.6	2763	24	ABA92252 Mouse Her-2/neu ex
2	144	90.6	2781	24	ABA92253 Her-2/neu extracel
3	84	52.8	2385	18	AA727235 Her2-GM-CSF immuno
4	80	50.3	69	21	AAA14815 DNA encoding a wil
5	80	50.3	201	21	AA62125 Her2 normal form t
6	80	50.3	2125	19	AAV21727 Humanised vector p
7	80	50.3	2871	21	AA250586 DC8scFv-erbB2EC fu
8	80	50.3	3600	21	AA889736 Human HER-2/neu co
9	80	50.3	3678	24	ABR66207 cDNA encoding huma
10	80	50.3	3768	17	AA740739 HER-2/neu oncogene
11	80	50.3	3768	20	AA301912 Human HER-2/neu on
12	80	50.3	3768	21	AAA09455 Human heregulin 2
13	80	50.3	3768	22	AAH23392 Human HER-2/neu pr
14	80	50.3	3768	24	AB235744 Human ERBB2 polynu
15	80	50.3	3768	24	ABX09987 Human ERBB2 DNA fr
16	80	50.3	3768	24	AA43935 Human HER-2 cDNA.
17	80	50.3	3768	24	AA43986 Human Her2 antigen
18	80	50.3	3768	24	ABV78168 Human ERBB2 DNA SE
19	80	50.3	3768	24	AA32743 Human Her-2/neu pr
20	80	50.3	3768	24	AA92250 Human Her-2/neu CD
21	80	50.3	3768	24	ABK10730 Human Her-2/neu DN
22	80	50.3	3768	24	ABL91709 Human polynucleoti
23	80	50.3	3768	24	ABK14058 Human HER2 (ErbB2)
24	80	50.3	4299	14	AAQ46083 Sequence encoding
25	80	50.3	4472	21	AAA14812 cDNA encoding the
26	80	50.3	4473	19	ABQ76220 Human tumour anti
27	80	50.3	4473	20	AA331071 HER-2 nucleic acid
28	80	50.3	4473	24	AB34969 Human gene express
29	80	50.3	4473	24	AA338904 Human Her-2 DNA.
30	80	50.3	4530	16	AA701585 Her-2/neu (ERBB2/c
31	80	50.3	4530	21	AA771253 Human HER2 gene.
32	80	50.3	4530	21	AA260815 Nucleotide sequenc
33	80	50.3	4530	22	AA197311 Human tyrosine kin
34	80	50.3	4530	24	AB335012 Human gene express
35	80	50.3	4530	24	ABV94128 Breast carcinoma r
36	80	50.3	4530	24	ABN85585 Human HER2-neu SEQ
37	80	50.3	4530	24	ABK83918 Human cDNA differe
38	80	50.3	4530	25	ACC50139 Breast cancer asso
39	80	50.3	4530	25	ABQ83856 Human Her2/Neu enc
40	80	50.3	9274	22	AA824297 HER2 transgene pla
41	80	50.3	9274	24	AA43934 Her-2 transgene pl
42	80	50.3	9274	24	ABK14057 Human HER2 (ErbB2)
43	79	49.7	1115	21	AA21778 Human breast and o
44	79	49.7	1755	24	AA32746 Human cDNA for the
45	79	49.7	1767	24	AA32744 Human cDNA for the

ALIGNMENTS

RESULT 1  
ABA92252  
ID ABA92252 standard; cDNA; 2763 BP.  
XX  
AC ABA92252;  
XX  
DT 17-JUN-2002 (first entry)  
XX  
DE Mouse Her-2/neu extracellular-phosphorylation domain fusion cDNA.  
XX  
KW Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;  
KW receptor; mouse; gene therapy; gene; ss.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2763









DE Humanised vector pITL-hHER/neu.  
 XX Vector: vaccine; tumour; antigen; plasmid pITL-hHER/neu;  
 KW human; HER-2/neu; C-erbB-2; breast cancer; ds.  
 XX

OS Chimeric - Homo sapiens.  
 OS Chimeric - Escherichia coli.  
 OS Synthetic.

XX Key Location/Qualifiers  
 FT CDS 13..921  
 FT /tag= a  
 FT /product= human HER-2/neu  
 FT polyA\_site 922..1181  
 FT /tag= b  
 FT /note= "combined splice and polyA sequences"  
 FT CDS 1195..1401  
 FT /tag= c  
 FT /note= "SupF gene"  
 FT misc\_feature 1412..1864  
 FT /tag= d  
 FT /note= "ColE1 origin of replication"  
 FT promoter 1877..2125  
 FT /tag= e  
 FT /note= "RANTES promoter"

PN W09806863-Al.  
 XX  
 XX PD 19-FEB-1998.  
 XX PF 14-AUG-1997; 97WO-US14306.  
 XX PR 14-AUG-1996; 96US-0023931.  
 XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX PI Nelson EL, Nelson PJ;  
 XX DR WPI; 1998-159552/14.  
 XX

XX Humanised polynucleotide vectors - comprising human derived promoter  
 PT and sequence acceptance site, used for the production of vaccines  
 XX  
 XX Example 7; Page 41-42; 125pp; English.

XX Plasmid pITL-hHER2/neu comprises base vector pITL (see AAV21724) and  
 CC a human HER-2/neu nucleic sequence. pITL-hHER2/neu was used to  
 CC evaluate the toxicity of anti-tumour vaccination in rats, and in  
 CC phase I and phase II trials to evaluate polynucleotide vaccination  
 CC in advanced breast cancer. Novel humanised vectors, which can be  
 CC based on pITL, comprise a human-derived promoter or mammalian  
 CC homologue which is functional in mammalian target tissue and cells  
 CC and a sequence acceptance site which accepts cDNA products from  
 CC RT-PCR cloning. The vectors are non-replicating in mammalian cells  
 CC but are capable of extended stable expression of the target  
 CC sequence, generating an immune response in immunised individuals.  
 CC The vectors selectively elicit immune responses to the target  
 CC sequences with little or no immune response to the other components  
 CC of the vectors.

XX Sequence 2125 BP; 449 A; 650 C; 598 G; 428 T; 0 other;

Alignment Scores:  
 Pred. No.: 0.124 Length: 2125  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.31% Indels: 0  
 DB: 19 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x AAV21727 (1-2125)

QY 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15

Db 112 GACAAGGCGTCCCGCGAGCAGAGAGCCAGCCCTCTGACGTC 156  
 RESULT 7  
 AAZ50586  
 ID AAZ50586 standard; DNA; 2871 BP.  
 XX  
 AC AAZ50586;  
 XX  
 DT 23-MAY-2000 (first entry)  
 XX  
 DE DC8scFv-erbB2EC fusion construct containing tetramerisation domain.  
 KW DC8 scFv; single-chain variable fragment; erbB2EC; extracellular domain;  
 KW human; fusion construct; tetramerisation domain; constant domain;  
 KW heteromultimer; multifunctional compound; melanoma; sarcoma;  
 KW immunoglobulin; cytostatic; immunostimulatory; antileukaemia; diagnosis;  
 KW antiproliferative; prevention; treatment; malignant; haematopoietic cell;  
 KW lymphoma; leukaemia; solid tumour; carcinoma; ds.  
 XX  
 OS Chimeric - Unidentified.  
 OS Chimeric - Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 10..2865  
 FT /tag= a  
 FT /product= "DC8scFv-erbB2EC fusion protein"  
 FT sig\_peptide 10..66  
 FT /tag= b  
 FT /label= Leader\_sequence  
 FT misc\_feature 67..390  
 FT /tag= c  
 FT /label= DC8scFv\_light\_chain\_variable\_region  
 FT misc\_feature 391..435  
 FT /tag= d  
 FT /label= Glycine-Serine-linker\_DNA  
 FT misc\_feature 436..771  
 FT /tag= e  
 FT /label= DC8scFv\_heavy\_chain\_variable\_region  
 FT misc\_feature 775..807  
 FT /tag= f  
 FT /note= "5' end of human IgG3 upper hinge region  
 FT with 3 additional nucleotides"  
 FT misc\_feature 808..924  
 FT /tag= g  
 FT /label= Human\_p53\_tetramerisation\_domain  
 FT misc\_feature 925..945  
 FT /tag= h  
 FT /label= Short\_peptide\_linker  
 FT misc\_feature 946..2844  
 FT /tag= i  
 FT /label= erbB2EC\_domain  
 FT misc\_feature 2845..2862  
 FT /tag= j  
 FT /label= His\_tag  
 XX WO200006605-A2.  
 XX  
 XX PD 10-FEB-2000.  
 XX  
 XX PF 28-JUL-1999; 99WO-EP05416.  
 XX  
 XX PR 28-JUL-1998; 98EP-0114082.  
 XX (MTCR-) MICROMET GES BIOMEDIZINISCHE FORSCHUNG.  
 XX Kufer P, Dreier T, Baeuerle PA, Borschert K, Zettl F;  
 XX WPI: 2000-195265/17.  
 XX P-PSDB: AAY44993.  
 XX New multifunctional compounds useful for preventing and/or treating  
 PT malignant cell growth and for detection and diagnosis

XX

PS Example 9; Fig 49; 166pp; English.

XX The patent discloses heteroninibodies which are multifunctional compounds  
 CC produceable in a mammalian host cell as a secretable and fully functional  
 CC heterodimer of two polypeptide chains, where one of the polypeptide  
 CC chains comprises, a C1-domain (constant domain of an immunoglobulin  
 CC heavy chain) and the other chain comprises C1-domain (constant domain of  
 CC an immunoglobulin light chain). The polypeptide chains further comprise,  
 CC fused to the constant domains at least two (polypeptide) chains having  
 CC different receptor or ligand functions, where further at least two of the  
 CC different (polypeptide) chains lack an intrinsic affinity for one another and  
 CC are linked via the constant domains. The heteroninibodies have  
 CC cytostatic, immunostimulatory, antileukaemia and antiproliferative  
 CC activities. These compounds can be used for diagnosing, preventing and  
 CC treating malignant cell growth related to malignancies of haematopoietic  
 CC cells e.g. lymphomas and leukaemias, or to solid tumours e.g. carcinomas,  
 CC melanomas and sarcomas.  
 CC The present sequence is a fusion construct comprising DC8  
 CC single-chain Fv (scFv) fragment at the N-terminus, extracellular  
 CC domain of human erbB2 at the C-terminus and a tetramerisation  
 CC domain between them. This construct was prepared to find out whether  
 CC an oligomerisation domain characterised in bacterial expression system  
 CC is applicable for expression of fully functional and secretable  
 CC recombinant protein in mammalian host cells. This tetrameric construct  
 CC was not expressed as secretable and fully functional protein  
 CC in mammalian cells. Hence general applicability of the tetramerisation  
 CC domain for oligomerisation strategies in mammalian cells was ruled out.

XX SQ Sequence 2871 BP; 598 A; 868 C; 834 G; 571 T; 0 other;

# Alignment Scores:

Pred. No.: 0.179 Length: 2871  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.31% Indels: 0  
 DB: 21 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x AAZ50586 (1-2871)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

DB 2797 GACAGGGCTGCCCGCCGAGCAGAGAGCCGCTCTGACGTCC 2841

## RESULT 8

AAA89736

ID AAA89736 standard; DNA; 3600 BP.

XX AAA89736;

XX 12-JAN-2001 (first entry)

XX Human HER-2/neu coding sequence.

DE Human HER-2/neu coding sequence.  
 XX Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;  
 KW breast cancer; prostate cancer; ovarian cancer; lung cancer;  
 KW colon cancer; ds.  
 XX Homo sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers

FF CDS 1..3600

FT /\*tag- a

FT /product= "HER-2/neu protein"

XX

PN WO200044899-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US02164.

XX 29-JAN-1999; 99US-0117976.

XX

PR

XX

XX

XX

XX

XX

XX

XX

PA (CORI-) CORIXA CORP.

XX (SMIK ) SMITHKLINE BEECHAM.

XX Cheever MA, Gheysen D;

XX WPI; 2000-505976/45.

DR P-PSDB; AAB21198, AAB21208.

XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins

PT useful for vaccinating against breast, ovarian, colon, lung and

PT prostate cancers -

XX Disclosure; Fig 15; 128pp; English.

XX The present sequence encodes the human HER-2/neu protein. HER-2/neu is

CC a member of the tyrosine kinase family of receptor-like glycoproteins

CC and shows homology to the epidermal growth factor receptor (EGFR). It

CC probably plays a part in cell growth and/or differentiation. The

CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising

CC a HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation

CC domain may be used to treat or prevent cancer by eliciting or

CC enhancing an immune response to the HER-2/neu protein. It may be used

CC to treat malignancies such as breast, ovarian, colon, lung and

CC prostate cancers, and may be used as an antigen to vaccinate against

CC these neoplasias.

XX SQ Sequence 3600 BP; 723 A; 1108 C; 1075 G; 694 T; 0 other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX







CC (ds) structure of not more than 49 consecutive nucleotides (nt), where  
CC at least a segment of one strand of the ds structure is complementary  
CC with the target gene and the cells are treated with interferon before  
CC introduction of dsRNA. The method is used to inhibit expression of  
CC target genes, particularly oncogenes, cytokine genes, id (not defined)  
CC protein genes; developmental or prion genes, or genes expressed in  
CC pathogenic organisms (particularly plasmodia) or in viruses or viroids  
CC (pathogenic in humans, animals or plants). Treating the cells with  
CC interferon greatly increases the extent to which dsRNA can inhibit  
CC expression of the target genes, and the effect is even greater when dsRNA  
CC are modified to increase their stability. ABX09936-ABX10075 represent  
CC gene fragments used to illustrate the method of the invention.

xx  
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:  
Pred. No.: 0.248 Length: 3768  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 24 Gaps: 0

JUNC\_SEQ3\_SEQ5 (1-30) x ABX09987 (1-3768)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
DB 1915 GACAAAGGGCTGCCCGCCCGAGCAGAGAGCCGCCCTCTGACGTCC 1959

Search completed: October 15, 2003, 23:48:41  
Job time : 31.4941 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:09:13 ; Search time 432.026 seconds  
(without alignments)  
2840.777 Million cell updates/sec

Title: JUNC\_SEQ3\_SEQ5

Perfect score: 159

Sequence: 1 DKGPABQBRASPLTSQNEGLGPASPLDSTF 30

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlip  
-Q=/cgn2\_1/USPTO\_Spool\_P/HOLLERAN480/runat\_15102003\_131912\_20536/app\_query.fasta\_1.4685  
-DB=GenEmbl -QMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=HOLLERAN480\_QCGN\_1\_1\_22724@runat\_15102003\_131912\_20536 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*

- 29: em\_vl.\*
- 30: em\_hgt\_hum.\*
- 31: em\_hgt\_inv.\*
- 32: em\_hgt\_other.\*
- 33: em\_hgt\_mus.\*
- 34: em\_hgt\_pln.\*
- 35: em\_hgt\_rod.\*
- 36: em\_hgt\_mam.\*
- 37: em\_hgt\_vrt.\*
- 38: em\_sy.\*
- 39: em\_hgtgo\_hum.\*
- 40: em\_hgtgo\_mus.\*
- 41: em\_hgtgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	144	90.6	2763	6	AX380942 Sequence
2	144	90.6	2781	6	AX380944 Sequence
3	85.5	53.8	186134	10	AL591390 Mouse DNA
4	85.5	53.8	214019	2	AC064803 Mus muscu
5	85	53.5	13450	9	AB096612 Homo sapi
6	85	53.5	16572	9	AB096614 Homo sapi
7	85	53.5	20271	9	AB096613 Homo sapi
8	85	53.5	30837	9	AY208911 Homo sapi
9	85	53.5	161815	9	AC079199 Homo sapi
10	85	53.5	168585	9	AC040933 Homo sapi
11	85	53.5	198008	2	AC142197 Homo sapi
12	84	52.8	2385	6	AR082744 Sequence
13	84	52.8	2385	6	AR099963 Sequence
14	84	52.8	2385	6	AR143949 Sequence
15	80	50.3	153	6	AX022984 Sequence
16	80	50.3	201	6	AX229723 Sequence
17	80	50.3	2871	6	AX023363 Sequence
18	80	50.3	3678	6	AX050514 Sequence
19	80	50.3	3768	6	AR034479 Sequence
20	80	50.3	3768	6	AX060704 Sequence
21	80	50.3	3768	6	AX201817 Sequence
22	80	50.3	3768	6	AX380923 Sequence
23	80	50.3	3768	6	AX384604 Sequence
24	80	50.3	3768	6	AX465456 Sequence
25	80	50.3	3768	6	AX467229 Sequence
26	80	50.3	3768	6	AX481438 Sequence
27	80	50.3	4473	6	AR080259 Sequence
28	80	50.3	4473	6	AR167390 Sequence
29	80	50.3	4473	9	HSEB2R
30	80	50.3	4530	6	AR202597 Sequence
31	80	50.3	4530	6	AR283481 Sequence
32	80	50.3	4530	6	AX282577 Sequence
33	80	50.3	4530	6	AX587649 Sequence
34	80	50.3	4530	6	AX644071 Sequence
35	80	50.3	4530	6	BD005474 Cellular
36	80	50.3	4530	6	I21124 Sequence 9
37	80	50.3	4530	6	I59745 Sequence 9
38	80	50.3	4530	9	HUMHER2A
39	80	50.3	9274	6	AX060703 Sequence
40	79	49.7	1755	6	AX384609 Sequence
41	79	49.7	1767	6	AX384607 Sequence
42	79	49.7	1773	6	AX384610 Sequence
43	79	49.7	1806	6	AX384608 Sequence
44	79	49.7	3780	4	AB008451 Canis fam
45	76	47.8	4062	10	D16295 Mesocricetu

ALIGNMENTS

RESULT 1

```
AX380942
LOCUS AX380942 2763 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 28 from Patent WO0212341.
ACCESSION AX380942
VERSION AX380942.1 GI:19575786
KEYWORDS
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
1 Cheever,M.A. and Gheysen,D.
AUTHORS Her-2/neu fusion proteins
TITLE Patent: WO 0212341-A 28 14-FEB-2002;
JOURNAL CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
source
1..2763
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mouse ECD-PD fusion protein cDNA"
BASE COUNT 571 a 855 c 772 g 565 t
ORIGIN
Alignment Scores:
Pred. No.: 8-97e-10 Length: 2763
Score: 144.00 Matches: 25
Percent Similarity: 100.00% Conservative: 5
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 90.57% Indels: 0
DB: 6 Gaps: 0

JUNC_SEQ3_SEQ5 (1-30) x AX380942 (1-2763)
QY 1 ASPLVSGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCACAGCAGACGAGCCAGCCAGCTCTCAGACGAGGACTTA 1977
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 1978 GGCCCTCCAGCCCATGGACAGCACCTTC 2007

RESULT 2
AX380944
LOCUS AX380944 2781 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 30 from Patent WO0212341.
ACCESSION AX380944
VERSION AX380944.1 GI:19575787
KEYWORDS
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
1 Cheever,M.A. and Gheysen,D.
AUTHORS Her-2/neu fusion proteins
TITLE Patent: WO 0212341-A 30 14-FEB-2002;
JOURNAL CORIXA CORPORATION (US) ; SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES
source
1..2781
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="mouse ECD-PD-Tcp0 fusion protein cDNA"
BASE COUNT 574 a 859 c 779 g 569 t
ORIGIN
Alignment Scores:
Pred. No.: 9.03e-10 Length: 2781
Score: 144.00 Matches: 25
Percent Similarity: 100.00% Conservative: 5
Best Local Similarity: 83.33% Mismatches: 0
Query Match: 90.57% Indels: 0
DB: 6 Gaps: 0
```

```
JUNC_SEQ3_SEQ5 (1-30) x AX380944 (1-2781)
QY 1 ASPLVSGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCACAGCAGACGAGCCAGCCAGCTCTCAGACGAGGACTTA 1977
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 1978 GGCCCTCCAGCCCATGGACAGCACCTTC 2007

RESULT 3
AL591390/c
LOCUS AL591390 186134 bp DNA linear ROD 17-NOV-2001
DEFINITION Mouse DNA sequence from clone RP23-355L10 on chromosome 11,
complete sequence.
ACCESSION AL591390
VERSION AL591390.8 GI:17017767
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 186134)
JOURNAL Direct Submission
COMMENT Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Nov 20, 2001 this sequence version replaced gi:16555512.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone
RP23-355L10. It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true right end of clone RP23-355L10 is at 186134 in this
sequence. The true right end of clone RP23-438D7 is at 2000 in this
sequence.

FEATURES
Location/Qualifiers
1..186134
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="11"
/clone="RP23-355L10"
/clone_lib="RPCI-23"
17991..18044
/note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."
misc_feature
47794..47812
/note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."
BASE COUNT 46287 a 47900 c 47689 g 44258 t
ORIGIN
```



## Alignment Scores:

```

Pred. No.:      5.11      Length:      186134
Score:          85.50      Matches:      18
Percent Similarity: 72.41%      Conservative: 3
Best Local Similarity: 62.07%      Mismatches: 5
Query Match:      53.77%      Indels:      3
DB:              10          Gaps:      1

```

JUNC\_SEQ3\_SEQ5 (1-30) x AL591390 (1-186134)

```

Qy 2 LysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGly 21
    :::::|||||
Db 70080 GAGGCTGT-----CACCTCGTGTCTCCCTTCTCTAGACGAGGACTTAGGC 70030

Qy 22 ProAlaSerProLeuAspSerThrPhe 30
    |||:::|||||:::|||||
Db 70029 CCCTCCAGCCCCATGGAGACGACCTTC 70003

RESULT 4
AC064803/c
LOCUS      AC064803          214019 bp      DNA      linear      HTG 02-NOV-2001
DEFINITION Mus musculus chromosome 11 clone RP23-16G14 map 11, WORKING DRAFT
SEQUENCE   SEQUENCE, 14 unordered pieces.
AC064803
VERSION    AC064803.4      GI:16596607
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 214019)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckheiser,B., Brown,A., Burkett,G.,
Campiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Melgrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 2, 2001 this sequence version replaced gi:12061538.
All repeats were identified using RepeatMasker:
Snit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9173
Center clone name: 16.G.14
----- Summary Statistics

```

```

Sequencing vector: M13; M77815; 54% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 208571 bases at least Q40
Consensus quality: 210746 bases at least Q30
Consensus quality: 211706 bases at least Q20
Insert size: 197000; agarose-fp
Insert size: 212719; sum-of-contigs
Quality coverage: 8.5 in Q20 bases; agarose-fp
Quality coverage: 7.8 in Q20 ba.

```

\* NOTE: This is a 'working draft' sequence. It currently consists of 14 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

* 1 21450: contig of 21460 bp in length
* 21461 21560: gap of 100 bp
* 21561 21561 21560: contig of 3285 bp in length
* 21561 21561 21560: contig of 3285 bp in length
* 24846 24846 24845: gap of 100 bp
* 24846 24846 24845: gap of 100 bp
* 28115 28115 28115: contig of 3170 bp in length
* 28116 28116 28115: gap of 100 bp
* 28216 35345: contig of 7130 bp in length
* 35346 35346 35345: gap of 100 bp
* 35346 35346 35345: gap of 100 bp
* 35446 43994: contig of 8549 bp in length
* 43995 44094: gap of 100 bp
* 44095 83787: contig of 39693 bp in length
* 83788 83887: gap of 100 bp
* 83888 96462: contig of 12575 bp in length
* 96463 96562: gap of 100 bp
* 96563 108784: contig of 12222 bp in length
* 108785 108884: gap of 100 bp
* 108885 122875: contig of 13991 bp in length
* 122876 122976 139026: contig of 16051 bp in length
* 139027 139126: gap of 100 bp
* 139127 159729: contig of 20603 bp in length
* 159730 159829: gap of 100 bp
* 159830 178173: contig of 18344 bp in length
* 178174 178273: gap of 100 bp
* 178274 200970: contig of 22697 bp in length
* 200971 201070: gap of 100 bp
* 201071 214019: contig of 12949 bp in length.

```

## FEATURES

## source

```

1..214019
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/map="11"
/clone="RP23-16G14"
/clone_lib="RPCI-23 Female Mouse BAC"

```

## misc\_feature

```

1..21460
/note="assembly_fragment"
/clone_end:SP6
vector_side:left"

```

## misc\_feature

```

21561..24845
/note="assembly_fragment"

```

## misc\_feature

```

24946..28115
/note="assembly_fragment"

```

## misc\_feature

```

28216..35345
/note="assembly_fragment"

```

## misc\_feature

```

35446..43994
/note="assembly_fragment"

```

## misc\_feature

```

44095..83787
/note="assembly_fragment"

```

## misc\_feature

```

83888..96462
/note="assembly_fragment"

```

## misc\_feature

```

96563..108784
/note="assembly_fragment"

```

```

misc_feature 10885..122875
              /note="assembly_fragment"
misc_feature 122976..139026
              /note="assembly_fragment"
misc_feature 139127..159729
              /note="assembly_fragment"
misc_feature 159830..178173
              /note="assembly_fragment"
misc_feature 178274..200970
              /note="assembly_fragment"
misc_feature 201071..214019
              /note="assembly_fragment"
              clone_end:=T
              vector_side:=right
BASE COUNT 52809 a 53348 c 53260 g 53244 t 1358 others
ORIGIN

Alignment Scores:
Pred. No.:      5.89      Length:      214019
Score:          85.50     Matches:      18
Percent Similarity: 72.41%  Conservative: 3
Best Local Similarity: 62.07% Mismatches: 5
Query Match:      53.77%  Indels:      3
DB:              2      Gaps:      1

JUNC_SEQ3_SEQ5 (1-30) x AC064803 (1-214019)
QY 2 LysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGly 21
Db 53608 GAGGCTGT-----CACCTCGTGTCTCCCTTCCTTAGACGAGGACTTAGC 53558

QY 22 ProAlaSerProLeuAspSerThrPhe 30
Db 53557 CCTCCAGCCCATGACAGACCTTC 53531

RESULT 5
AB096612
LOCUS      AB096612      13450 bp      DNA      linear      PRI 21-MAY-2003
DEFINITION Homo sapiens DNA, 13kb-normal EcoRI sequence.
ACCESSION  AB096612
VERSION     AB096612.1  GI:30962525
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Kuwahara,Y., Tanabe,C., Ikeuchi,T., Aoyagi,K., Nishigaki,M.,
            Sakamoto,H., Hoshinaga,K., Yoshida,T., Sasaki,H. and Terada,M.
            Presence of novel mechanisms other than the BFB cycles in
            amplification of human oncogene
JOURNAL     Unpublished
AUTHORS     Sasaki,H.
TITLE       Direct Submission
JOURNAL     Submitted (20-NOV-2002) Hiroki Sasaki, National Cancer Center
            Research Institute, Genetics Division; Tsukiji 5-1-1, Chuo-ku,
            Tokyo 104-0045, Japan (E-mail:hksasaki@gan2.res.ncc.go.jp,
            Tel:81-3-3542-2511(ex.4402), Fax:81-3-3541-2685)
FEATURES
source
1..13450
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="17"
   /map="17q12"
   /clone="13kb-normal EcoRI DNA"
684..2046
   /gene="c-ERBB-2"
684..936
   /gene="c-ERBB-2"
1078..2046
   /gene="c-ERBB-2"
1078..2046
   /gene="c-ERBB-2"
1078..2046
   /gene="c-ERBB-2"

misc_feature 10885..122875
              /note="assembly_fragment"
misc_feature 122976..139026
              /note="assembly_fragment"
misc_feature 139127..159729
              /note="assembly_fragment"
misc_feature 159830..178173
              /note="assembly_fragment"
misc_feature 178274..200970
              /note="assembly_fragment"
misc_feature 201071..214019
              /note="assembly_fragment"
              clone_end:=T
              vector_side:=right
BASE COUNT 52809 a 53348 c 53260 g 53244 t 1358 others
ORIGIN

Alignment Scores:
Pred. No.:      5.89      Length:      214019
Score:          85.50     Matches:      18
Percent Similarity: 72.41%  Conservative: 3
Best Local Similarity: 62.07% Mismatches: 5
Query Match:      53.77%  Indels:      3
DB:              2      Gaps:      1

JUNC_SEQ3_SEQ5 (1-30) x AC064803 (1-214019)
QY 2 LysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGly 21
Db 53608 GAGGCTGT-----CACCTCGTGTCTCCCTTCCTTAGACGAGGACTTAGC 53558

QY 22 ProAlaSerProLeuAspSerThrPhe 30
Db 53557 CCTCCAGCCCATGACAGACCTTC 53531

RESULT 5
AB096612
LOCUS      AB096612      13450 bp      DNA      linear      PRI 21-MAY-2003
DEFINITION Homo sapiens DNA, 13kb-normal EcoRI sequence.
ACCESSION  AB096612
VERSION     AB096612.1  GI:30962525
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Kuwahara,Y., Tanabe,C., Ikeuchi,T., Aoyagi,K., Nishigaki,M.,
            Sakamoto,H., Hoshinaga,K., Yoshida,T., Sasaki,H. and Terada,M.
            Presence of novel mechanisms other than the BFB cycles in
            amplification of human oncogene
JOURNAL     Unpublished
AUTHORS     Sasaki,H.
TITLE       Direct Submission
JOURNAL     Submitted (20-NOV-2002) Hiroki Sasaki, National Cancer Center
            Research Institute, Genetics Division; Tsukiji 5-1-1, Chuo-ku,
            Tokyo 104-0045, Japan (E-mail:hksasaki@gan2.res.ncc.go.jp,
            Tel:81-3-3542-2511(ex.4402), Fax:81-3-3541-2685)
FEATURES
source
1..13450
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="17"
   /map="17q12"
   /clone="13kb-normal EcoRI DNA"
684..2046
   /gene="c-ERBB-2"
684..936
   /gene="c-ERBB-2"
1078..2046
   /gene="c-ERBB-2"
1078..2046
   /gene="c-ERBB-2"

misc_feature 10885..122875
              /note="3' exon"
              11349..11517
              /gene="GRB7"
              11349..11517
              /gene="GRB7"
              11349..11517
              /gene="GRB7"
BASE COUNT 2748 a 3818 c 3387 g 3496 t 1 others
ORIGIN

Alignment Scores:
Pred. No.:      0.415     Length:      13450
Score:          85.00     Matches:      23
Percent Similarity: 46.00%  Conservative: 0
Best Local Similarity: 46.00% Mismatches: 3
Query Match:      53.46%  Indels:      24
DB:              9      Gaps:      2

JUNC_SEQ3_SEQ5 (1-30) x AB096612 (1-13450)
QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
Db 96 CCTGCAGAGGGTGGGAAGGAGAGATGAGTCCAGATATGCCAGGCCCTCAGCGAAGGCTGC 155
QY 16 -----GlnAsnGluAspLeu 20
Db 156 ATGCTGGGTGGGGAGGGGCCACCATCTCTCTCTCTCCACAGATGAGGACTTG 215
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 216 GGGCCAGCAGTCCCTTGGACAGCACCTTC 245

RESULT 6
AB096614
LOCUS      AB096614      16572 bp      DNA      linear      PRI 21-MAY-2003
DEFINITION Homo sapiens DNA, 16.5kb-T5 rearranged sequence.
ACCESSION  AB096614
VERSION     AB096614.1  GI:30962527
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Kuwahara,Y., Tanabe,C., Ikeuchi,T., Aoyagi,K., Nishigaki,M.,
            Sakamoto,H., Hoshinaga,K., Yoshida,T., Sasaki,H. and Terada,M.
            Presence of novel mechanisms other than the BFB cycles in
            amplification of human oncogene
JOURNAL     Unpublished
AUTHORS     Sasaki,H.
TITLE       Direct Submission
JOURNAL     Submitted (20-NOV-2002) Hiroki Sasaki, National Cancer Center
            Research Institute, Genetics Division; Tsukiji 5-1-1, Chuo-ku,
            Tokyo 104-0045, Japan (E-mail:hksasaki@gan2.res.ncc.go.jp,
            Tel:81-3-3542-2511(ex.4402), Fax:81-3-3541-2685)
FEATURES
source
1..16572
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="17"
   /map="17q12"
   /clone="16.5kb-T5 rearranged DNA"
684..2046
   /tissue.type="gastric cancer"
684..936
   /gene="c-ERBB-2"
1078..2046
   /gene="c-ERBB-2"
1078..2046
   /gene="c-ERBB-2"
11349..11517
   /note="3' exon"
11349..11517
   /gene="GRB7"
11349..11517
   /gene="GRB7"

```

```

misc_feature 11419
/ gene="GRB7"
/ note="inter-amplicon junction"
BASE COUNT 3438 a 4677 c 4063 g 4394 t
ORIGIN
Alignment Scores:
Pred. No.: 0.513 Length: 16572
Score: 85.00 Matches: 23
Percent Similarity: 46.00% Conservative: 0
Best Local Similarity: 46.00% Mismatches: 3
Query Match: 53.46% Indels: 24
DB: 9 Gaps: 2
JUNC_SEQ3_SEQ5 (1-30) x AB096614 (1-16572)
QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
Db 96 CCTGCAGAGGGTGGGAAGAGAGAGATGCTCCAGTATGCCAGGCCCTCACGGAAGGCTGC 155
QY 16 -----GlnAsnGluAspLeu 20
Db 156 ATGCTGGGCTGGGAGGGGCCACCATCTGCTCTCTCTCCACACAGATGAGGACTTG 215
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 216 GGGCCAGCAGTCCCTTGGACAGCACCTTC 245
RESULT 7
AB096613 20271 bp DNA linear PRI 21-MAY-2003
LOCUS Homo sapiens DNA, 20kb-MKN7 rearranged sequence.
DEFINITION AB096613
ACCESSION AB096613.1 GI:30962526
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kuwahara,Y., Tanabe,C., Ikeuchi,T., Aoyagi,K., Nishigaki,M.,
Sakamoto,H., Hoshinaga,K., Yoshida,T., Sasaki,H. and Terada,M.
TITLE Presence of novel mechanisms other than the BFB cycles in
amplification of human oncogene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 20271)
AUTHORS Sasaki,H.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-2002) Hiroki Sasaki, National Cancer Center
Research Institute, Genetics Division; Tsukiji 5-1-1, Chuo-ku,
Tokyo 104-0045, Japan (E-mail:hksasaki@egan2.res.ncc.go.jp,
Tel:81-3-3542-2511(ex.4402), Fax:81-3-3541-2685)
FEATURES
source
Location/Qualifiers
1..20271
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q12"
/clone="20kb-MKN7 rearranged DNA."
/cell_line="MKN7"
/tissue_type="gastric cancer"
684..2046
/ gene="c-ERBB-2"
684..936
/ exon
/ gene="c-ERBB-2"
1078..2046
/ exon
/ gene="c-ERBB-2"
/ note="3' exon"
11349..11419
/ gene
11349..11419
/ exon
/ gene="GRB7"
/ gene="GRB7"
misc_feature 11419
/ gene="GRB7"
/ note="inter-amplicon junction"
BASE COUNT 4372 a 5349 c 5309 g 5241 t
ORIGIN
Alignment Scores:
Pred. No.: 0.629 Length: 20271
Score: 85.00 Matches: 23
Percent Similarity: 46.00% Conservative: 0
Best Local Similarity: 46.00% Mismatches: 3
Query Match: 53.46% Indels: 24
DB: 9 Gaps: 2
JUNC_SEQ3_SEQ5 (1-30) x AB096613 (1-20271)
QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
Db 96 CCTGCAGAGGGTGGGAAGAGAGAGATGCTCCAGTATGCCAGGCCCTCACGGAAGGCTGC 155
QY 16 -----GlnAsnGluAspLeu 20
Db 156 ATGCTGGGCTGGGAGGGGCCACCATCTGCTCTCTCTCCACACAGATGAGGACTTG 215
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 216 GGGCCAGCAGTCCCTTGGACAGCACCTTC 245
RESULT 8
AY208911 30837 bp DNA linear PRI 12-JAN-2003
LOCUS Homo sapiens v-erb-b2 erythroblastic leukemia viral oncogene
homolog 2, neuro/glioblastoma derived oncogene homolog (avian)
(ERBB2) gene, complete cds.
DEFINITION AY208911
ACCESSION AY208911.1 GI:27657999
VERSION AY208911
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 30837)
AUTHORS Rieder,M.J., Livingston,R.J., Daniels,M.R., Montoya,M.A.,
Chung,M.-W., Miyamoto,K.E., Nguyen,C.P., Nguyen,D.A., Poel,C.L.,
Robertson,P.D., Schackwitz,W.S., Sherwood,J.K., Witrak,L.A. and
Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-2002) Genome Sciences, University of Washington,
1705 NE Pacific, Seattle, WA 98195, USA
COMMENT To cite this work please use: NIEHS-SNPs, Environmental Genome
Project, NIEHS ES15478, Department of Genome Sciences, Seattle, WA
(URL: http://egp.gs.washington.edu).
FEATURES
Location/Qualifiers
1..30837
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..573
/ note="Region not scanned for variation"
288..584
/ repeat_region
/ rpt_family="Alu"
/ rpt_type="dispersed"
691
/ variation
/ frequency="0.07"
/ replace="a"
727
/ variation
/ frequency="0.01"
/ replace="g"

```

```

variation
902 /frequency="0.01"
    /replace="a"
variation
975 /frequency="0.03"
    /replace="t"
gene
1253. .29789
    /gene="ERBB2"
mRNA
join(1253. .1421,8102. .8253,9440. .9653,10437. .10571,
10932. .11000,11205. .11320,11459. .11600,13047. .13166,
13441. .13567,16405. .16478,16585. .16855,16859. .17058,
17420. .17552,17634. .17724,18439. .18599,20908. .20955,
24439. .24577,24658. .24780,25032. .25130,25858. .26043,
26181. .26336,26459. .26534,26839. .26985,27694. .27791,
27947. .28135,28427. .28679,28821. .29789)
    /gene="ERBB2"
/product="v-erb-b2 erythroblastic leukemia viral oncogene
homolog 2, neuro/glioblastoma derived oncogene homolog
(avian)"
CDS
join(1349. .1421,8102. .8253,9440. .9653,10437. .10571,
10932. .11000,11205. .11320,11459. .11600,13047. .13166,
13441. .13567,16405. .16478,16585. .16855,16859. .17058,
17420. .17552,17634. .17724,18439. .18599,20908. .20955,
24439. .24577,24658. .24780,25032. .25130,25858. .26043,
26181. .26336,26459. .26534,26839. .26985,27694. .27791,
27947. .28135,28427. .28679,28821. .29176)
    /gene="ERBB2"
/codon_start=1
/product="v-erb-b2 erythroblastic leukemia viral oncogene
homolog 2, neuro/glioblastoma derived oncogene homolog
(avian)"
/db_xref="GI:27658000"
/translation="MELAALCRWGLLLALLPCCAASTOVCTGCTDMKLRLPASPETHLD
MLRHLYQGVQGNLELYLPTNASLSFLQDIQVQGYVLIAHNOVRVPLQRLKIV
RGTLDFALVLDNDGDLNNTPTVTGASPGGURELQRLSLTEILKGGVLIQRNPQ
LCYQDILFMKNQALTLIDTNRGRACPCPMCKSRGWGESSEDCQSLRT
VCAGGACRGKPLPTDCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNT
DTFESMPNPGRYTFCGSCVTACPYNLSTDVSCVLPVHNOEVTADGQVQCEKC
SKPCARVCYGLGMEHLREVRVTSANIOFAGCKIEGSLAFLPESFDGDPASNTAPL
QPEQLQFTLEITGTYLISAMPDLSPLVFQNLQVIRGRILHNGAYSLTQGLGI
SMGLSRELGSGLIHHNTHLCFVHTVPMWDLFRNPHQALLHTANPEDECVEG
LACHQACRGPGTQCVNCQSFQROECVECRVLQGLPREYVNAHGLPCHPCE
CQFQNGSVTCFGEADQVACAHYKDPFCVACPSGVKPDLSYMP IWKFPDEEACQ
PCPINTHSCVDLDKQCPAEQORASPLTIIISAVGILLVVLVGVFGILIKRROKI
RKYTMRLQETELVPLTPSGAMPNQAMRIKTELKRVKVLGSGAGTGVYKGIWI
PDGENVKIPVALIKVLENTSPKANKELIDEAIVMAGVSPYVSRLLGICLTSTVQLVT
QLMPYCLLDHVRNRRGSGQDLLNWCQIAGMSILEDVRLVHDLAARNVLYKSP
NHVKTDFGLARLDLIDETEHADGKVP IKMWALESILRRFTHQSDVMYSYGVYWE
LMTFGAKPYDIPAREIPDLLEKGERLPQPICTIDVYIMVCKWMIDSECRPREL
VSEFSRMARDPQRFVYIIONEDLGPASPLDTFVRSLLDDMDGLVDABEYLVPOQGF
FCPDPAAGAGVHRRHSSTRSGGDLTLGLEPSEEAAPRSLAPSGAGSDVFDG
DLGMNAKGLQSLPHDPSPLARYSEDVPLPSETDGYVAPLTSPQPEYVNPQDVR
PQPSPREGLPAAPAGATLERAKTLPSPKNGVVKDFAFGGAIVENPEYILTPQGGA
PQPHPPPAFSPAFDNLNYYWDQDPPERGAPPTFKGTPTAENPEYILGLDVPV"
1422. .1712
    /gene="ERBB2"
    /note="Region not scanned for variation"
variation
1839 /gene="ERBB2"
    /frequency="0.02"
    /replace="c"
misc_feature
1887. .6576
    /gene="ERBB2"
    /note="Region not scanned for variation"
repeat_region
2315. .2536
    /rpt_family="MIR"
repeat_region
3217. .3417
    /rpt_family="Alu"
    /rpt_type=dispersed
repeat_region
3837. .3974
    /rpt_family="MIR"
    /rpt_type=dispersed
    /gene="ERBB2"
    /frequency="0.02"
    /rpt_type=dispersed
4784. .4858
    /rpt_family="MIR"
    /rpt_type=dispersed
5404. .5693
    /rpt_family="Alu"
    /rpt_type=dispersed
6074. .6282
    /rpt_family="MIR"
    /rpt_type=dispersed
6690
    /gene="ERBB2"
    /replace="t"
6952. .6953
    /gene="ERBB2"
    /replace="g"
6957
    /frequency="0.02"
    /replace="g"
6960
    /gene="ERBB2"
    /replace="g"
7036
    /gene="ERBB2"
    /replace="t"
7146
    /gene="ERBB2"
    /replace="t"
7151
    /gene="ERBB2"
    /replace="t"
7825
    /gene="ERBB2"
    /replace="t"
7874
    /gene="ERBB2"
    /replace="a"
8127
    /gene="ERBB2"
    /replace="a"
8467. .8473
    /gene="ERBB2"
    /replace="a"
8574
    /gene="ERBB2"
    /replace="a"
8575. .8707
    /rpt_family="MIR"
    /rpt_type=dispersed
8836. .9130
    /rpt_family="Alu"
    /rpt_type=dispersed
8875. .9119
    /gene="ERBB2"
    /note="Region not scanned for variation"
9136. .9214
    /rpt_family="MIR"
    /rpt_type=dispersed
9595
    /gene="ERBB2"
    /frequency="0.02"

```

```

/replacement="g"
/region .9770
/rpt_family="tip100"
/rpt_type-dispersed
9766
/region="ERBB2"
/region="0.01"
/region="t"
9799 .9946
/rpt_family="tip100"
/rpt_type-dispersed
9949 .10151
/rpt_family="MIR"
/rpt_type-dispersed
10009
/region="ERBB2"

```

```

Alignment Scores:
Pred. No.: 0.963 Length: 30837
Score: 85.00 Matches: 23
Percent Similarity: 46.00% Conservative: 0
Best Local Similarity: 46.00% Mismatches: 3
Query Match: 53.46% Indels: 24
DB: 9 Gaps: 2

```

JUNC\_SEQ3\_SEQ5 (1-30) x AY208911 (1-30837)

```

QY 5 ProLAGlu-----GlnArgAlaSerProLeuThrSer----- 15
Db 27839 CTTGAGAGGCTGGGAGGAGAGATGACTCCAGTATGCCAGCCCTCCACGAGGCTGC 27898
QY 16 -----GlnAsnGluAspLeu 20
Db 27899 ATGTGGCTGGGGGGGCCACCATCTGCTCTCTCTCTCCACAGATGAGGACTTG 27958
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 27959 GGCCAGCCAGCTCCCTGGACAGCACCTTC 27988

```

```

RESULT 9
AC079199 AC079199 161815 bp DNA linear PRI 02-OCT-2002
LOCUS Homo sapiens chromosome 17, clone RP11-94L15, complete sequence.
DEFINITION AC079199
ACCESSION AC079199
VERSION AC079199.9 GI:23462913
KEYWORDS HTG
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 161815)
Birken, B., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 17, clone RP11-94L15
Unpublished
2 (bases 1 to 161815)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bedalov, F., Boguslavskiy, L.,
Boukhaltier, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., LaRocque, K.,
Lanazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G.,
Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,
McPherson, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Pierre, N., Pisanic, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,
Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,

```

```

Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (23-AUG-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 161815)
Birken, B., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhaltier, B.,
Barna, N., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Iliev, I., Johnson, R., Jones, C., Kann, L.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lieu, C., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (08-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 161815)
Birken, B., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhaltier, B.,
Barna, N., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Iliev, I., Johnson, R., Jones, C., Kann, L.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lieu, C., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (02-OCT-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 2, 2002 this sequence version replaced gi:22758800.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L10635
Center clone name: 94_L15
----- Location/Qualifiers
1. 161815
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-94L15"
/clone_lib="RP11-11 Human Male BAC"
7. 470
/note="PCR product sequence only"

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
FEATURES
source
unsure

```

```

repeat_region complement(262..317)
unsure /rpt_family="MIR"
repeat_region /note="PCR product sequence only"
repeat_region 1689..1986
repeat_region 5033..5179
repeat_region 6113..6233
repeat_region /rpt_family="MIR"
repeat_region 6418..6460
repeat_region /rpt_family="L3"
repeat_region 7910..7930
repeat_region /rpt_family="AT-rich"
repeat_region 8422..8476
repeat_region /rpt_family="(CCTG)n"
repeat_region 8513..8659
repeat_region /rpt_family="MIR"
repeat_region complement(8917..9066)
repeat_region /rpt_family="MIR3"
repeat_region 12588..12697
repeat_region /rpt_family="G-rich"
repeat_region complement(13513..13824)
repeat_region /rpt_family="AluSp"
repeat_region complement(13836..13937)
repeat_region /rpt_family="MIR"
repeat_region 18739..18939
repeat_region /rpt_family="MIR"
repeat_region 21075..21274
repeat_region /rpt_family="MIR"
repeat_region 21447..21714
repeat_region /rpt_family="MIR"
repeat_region complement(23385..23684)
repeat_region /rpt_family="AluSx"
repeat_region 24620..24819
repeat_region /rpt_family="MIR"
repeat_region 24991..25298
repeat_region /rpt_family="L2"
repeat_region complement(25299..25605)
repeat_region /rpt_family="AluJb"
repeat_region 25606..25686
repeat_region /rpt_family="L2"
repeat_region 25687..25838
repeat_region /rpt_family="MIR"
repeat_region 25839..25952
repeat_region /rpt_family="LTR54B"
repeat_region complement(25953..26241)
repeat_region /rpt_family="AluSp"
repeat_region 26242..26378
repeat_region /rpt_family="LTR54B"
repeat_region 26379..26695
repeat_region /rpt_family="AluY"
repeat_region 26696..26901
repeat_region /rpt_family="LTR54B"
repeat_region 26902..26985
repeat_region /rpt_family="MIR"
repeat_region 26986..27224
repeat_region /rpt_family="L2"
repeat_region 27243..27550
repeat_region /rpt_family="AluSg"
repeat_region 27695..28007
repeat_region /rpt_family="L2"
repeat_region 28290..28600
repeat_region /rpt_family="AluSx"
repeat_region 29462..29600
repeat_region /rpt_family="MIR"
repeat_region 29643..29967
repeat_region /rpt_family="AluY"
repeat_region complement(30700..31000)
repeat_region /rpt_family="AluSp"
repeat_region complement(31002..31248)
repeat_region /rpt_family="AluJb"
repeat_region 31521..31572

```

```

/rpt_family="L2"
31577..31796
/rpt_family="MIR"
repeat_region 32751..33072
repeat_region /rpt_family="L2"
33181..33293
/rpt_family="L2"
34072..34376
/rpt_family="AluSp"
complement(34542..34833)
/rpt_family="AluJb"
complement(35218..35297)
/rpt_family="L2"
35320..35431
/rpt_family="AluSp"
35433..35500
/rpt_family="AluSp/q"
complement(35581..35702)
/rpt_family="L2"
35710..36008
/rpt_family="AluSc"
36009..36023

Alignment Scores:
Pred. No.: 5.18 Length: 161815
Score: 85.00 Matches: 23
Percent Similarity: 46.00% Conservative: 0
Best Local Similarity: 46.00% Mismatches: 3
Query Match: 53.46% Indels: 24
DB: 9 Gaps: 2

JUNC_SEQ3_SEQ5 (1-30) x AC079199 (1-161815)
QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
Db 71166 CCTGCAGAGGGTGGGAAGGAGAGATGATGCCAGGCCCTCCACGGAGGCTGC 71225
QY 16 -----GlnAsnGluAspLeu 20
Db 71226 ATGCTGGGCTGGGAGGGCCACCATCTGCTCTCTCTCCACAGATGAGGACTTG 71285

RESULT 10
AC040933 AC040933 168585 bp DNA linear PRI 01-OCT-2002
LOCUS Homo sapiens chromosome 17, clone CTD-2019C10, complete sequence.
DEFINITION AC040933
ACCESSION AC040933
VERSION AC040933.12 GI:23396303
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 168585)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone CTD-2019C10
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 168585)
AUTHORS Birren,B., Linton,L., Barna,N., Bastien,V., Beda,F.,
Anderson,S., Baldwin,J., Boukhalter,B., Brown,A., Burkett,G.,
Boguslavsky,L., Borkhater,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,

```

Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

#### TITLE

#### JOURNAL

#### REFERENCE

#### AUTHORS

Submitted (11-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 168585)  
 Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

#### Direct Submission

Submitted (01-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 4 (bases 1 to 168585)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

#### Direct Submission

Submitted (01-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Oct 1, 2002 this sequence version replaced gi:22597589.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

#### Center

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L8894

Center clone name: 2019\_C\_10

-----

Only the last 167.6 kb of this clone are being submitted.

The remainder overlaps accession number AC097491 [WICGR project L11967].

#### FEATURES

#### source

Location/Qualifiers  
 1. 168585  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="17"  
 /map="17"  
 /clone="CTD-2019C10"  
 /clone\_lib="CITD1 Human BAC"  
 /complement(1359..1516)  
 /rpt\_family="MIR"  
 /complement(1538..1754)  
 /rpt\_family="MIR"  
 2117..2419  
 /rpt\_family="AluSq"  
 /complement(2609..2748)  
 /rpt\_family="MIR"  
 4091..4278  
 /rpt\_family="MIR"  
 /complement(4334..4339)  
 /note="<30 qual SNGL region"  
 4409..4725  
 /rpt\_family="AluJb"  
 4782..4942  
 /rpt\_family="L1ME1"  
 4986..5046  
 /rpt\_family="(TGAA)n"  
 /complement(5492..5657)  
 /rpt\_family="MIR"  
 6438..6626  
 /rpt\_family="CT-rich"  
 8771..8827  
 /rpt\_family="CA-rich"  
 10677..10875  
 /rpt\_family="MIR3"  
 /complement(12314..12419)  
 /rpt\_family="L2"  
 12443..12588  
 /rpt\_family="AluJo"  
 12589..12885  
 /rpt\_family="AluSp"  
 12886..13057  
 /rpt\_family="AluJo"  
 /complement(14521..14553)  
 /note="<30 qual SNGL region"  
 /complement(14571..14592)  
 /note="<30 qual SNGL region"  
 /complement(14609..14614)  
 /note="<30 qual SNGL region"  
 /complement(14624..14629)  
 /note="<30 qual SNGL region"  
 /complement(14639..14666)  
 /note="<30 qual SNGL region"  
 15796..15849  
 /rpt\_family="GA-rich"  
 16715..17025  
 /rpt\_family="AluSx"  
 /complement(17213..17283)  
 /rpt\_family="MIR"  
 /complement(17647..17651)  
 /note="<30 qual SNGL region"  
 19655..20204  
 /rpt\_family="L2"  
 20443..20649  
 /rpt\_family="MIR"  
 20663..20928  
 /rpt\_family="AluJo"  
 21295..21301  
 /note="<30 qual SNGL region"  
 21587..21929  
 /rpt\_family="L2"  
 /complement(22041..22147)  
 /rpt\_family="MIR"

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 234.794 Seconds  
(without alignments)  
3105.423 Million cell updates/sec

Title: JUNC\_SEQ3\_SEQ4  
Perfect score: 159  
Sequence: 1 DRGCPAEQRAPLTSQDNEDGPASPLDSTF 30

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framet\_p2n.model -DEV=xlp  
-Q=/cgp2\_1/USPTO.spool\_p/HOLLERAN480/runat\_15102003\_131913\_20548/app\_query.fasta\_1.4685  
-DB=EST -QPMI=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFM=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480@cgn.1.12645 @runat\_15102003\_131913\_20548 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*\*  
1: em.estba:\*\*  
2: em.esthum:\*\*  
3: em.estin:\*\*  
4: em.estnu:\*\*  
5: em.estov:\*\*  
6: em.estpl:\*\*  
7: em.estro:\*\*  
8: em.htc:\*\*  
9: gb.est1:\*\*  
10: gb.est2:\*\*  
11: gb.htc:\*\*  
12: gb.est3:\*\*  
13: gb.est4:\*\*  
14: gb.est5:\*\*  
15: em.estfun:\*\*  
16: em.estom:\*\*  
17: em.gss\_hum:\*\*  
18: em.gss\_inv:\*\*  
19: em.gss\_pln:\*\*  
20: em.gss\_vrt:\*\*  
21: em.gss\_fun:\*\*  
22: em.gss\_mam:\*\*  
23: em.gss\_mus:\*\*  
24: em.gss\_pro:\*\*  
25: em.gss\_rod:\*\*  
26: em.gss\_phg:\*\*  
27: em.gss\_vrl:\*\*  
28: gb.gss1:\*\*

29: gb.gss2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	85.5	53.8	618	28	AZ699313	AZ699313 RPCI-23-2
2	85	53.5	422	9	AW606421	AW606421 RC3-HT037
3	85	53.5	439	10	BE157169	BE157169 RC3-HT037
4	85	53.5	944	14	CD558559	CD558559 AGENCOURT
5	81	50.9	458	10	BE157103	BE157103 RC3-HT037
6	80	50.3	378	10	BF869854	BF869854 IL3-ET011
7	80	50.3	411	2	HS0069572	Bx479682 Homo sapi
8	80	50.3	433	10	BF762317	BF762317 IL2-CS004
9	80	50.3	450	12	BG988031	BG988031 MR2-HT116
10	80	50.3	466	12	BM720186	BM720186 UT-E-E00-
11	80	50.3	546	12	BI194790	BI194790 602948150
12	80	50.3	547	9	AA443351	AA443351 zW85b05.r
13	80	50.3	561	12	BM720098	BM720098 UT-E-E00-
14	80	50.3	607	12	BM083444	BM083444 imageqc_6
15	80	50.3	613	9	AW410534	AW410534 fh06h06.x
16	80	50.3	666	13	BX113765	BX113765 BX113765
17	80	50.3	667	12	BM048624	BM048624 603623984
18	80	50.3	674	12	BM758528	BM758528 K-EST0038
19	80	50.3	741	10	BE615590	BE615590 601278973
20	80	50.3	881	9	AI906106	AI906106 RC-BT105-
21	80	50.3	886	13	BQ882315	BQ882315 AGENCOURT
22	80	50.3	893	14	CA487631	CA487631 AGENCOURT
23	80	50.3	905	13	BU541491	BU541491 AGENCOURT
24	80	50.3	909	10	BG289457	BG289457 602381495
25	80	50.3	922	14	CA453990	CA453990 AGENCOURT
26	80	50.3	1009	13	BQ647343	BQ647343 AGENCOURT
27	79	49.7	278	10	BF923905	BF923905 QV4-NT025
28	79	49.7	294	10	BF757857	BF757857 CM4-CT057
29	79	49.7	337	10	BE091682	BE091682 IL2-BT073
30	79	49.7	348	9	AI909847	AI909847 QV-BT225-
31	79	49.7	378	14	CB266492	CB266492 1005398 H
32	79	49.7	423	12	BM703962	BM703962 UT-E-CK1-
33	79	49.7	464	2	HS0067221	Bx477782 Homo sapi
34	79	49.7	496	12	BM795191	BM795191 K-EST0076
35	79	49.7	505	12	BM854421	BM854421 K-EST0136
36	79	49.7	531	12	BM830170	BM830170 K-EST0103
37	79	49.7	536	14	CB129403	CB129403 K-EST0179
38	79	49.7	547	12	BM787824	BM787824 K-EST0066
39	79	49.7	567	2	HS0077766	Bx487579 Homo sapi
40	79	49.7	574	12	BM829991	BM829991 K-EST0103
41	79	49.7	630	2	HS0073298	Bx483130 Homo sapi
42	79	49.7	691	14	CB853376	CB853376 UT-CF-FN0
43	79	49.7	767	10	BF240297	BF240297 601905830
44	79	49.7	849	14	CA489799	CA489799 AGENCOURT
45	79	49.7	852	14	CD516283	CD516283 AGENCOURT

ALIGNMENTS

RESULT 1  
AZ699313  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE

AZ699313  
RPCI-23-240N16.TJ RPCI-23 Mus musculus genomic clone  
, genomic survey sequence.  
AZ699313  
GI:12419623  
GSS.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 618)

618 bp  
DNA  
linear  
GSS 24-JAN-2001

**AUTHORS** Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P., and Fraser, C.M.

**TITLE** Mouse BAC End Sequences from Library RPCI-23

**JOURNAL COMMENT** Unpublished  
Other\_GSSs: RPCI-23-240N16.TV  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 240 row: N column: 16  
Seq primer: SP6  
Class: BAC ends.

# FEATURES

source  
1. .618  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-23-240N16"  
/sex="Female"  
/lab\_host="DH10B"  
/clone\_lib="RPCI-23"  
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site.1: EcoRI; Site.2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 110 a 214 c 151 g 143 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.347 Length: 618  
Score: 85.50 Matches: 18  
Percent Similarity: 72.41% Conservative: 3  
Best Local Similarity: 62.07% Mismatches: 5  
Query Match: 53.77% Indels: 3  
DB: 28 Gaps: 1

JUNC\_SEQ3\_SEQ4 (1-30) x AZ699313 (1-618)

QY 2 LysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGly 21  
Db 165 GAGGCTGT-----CACCTCGTGTCTCCCTTCTCTCCAGAGGACTAGGC 215

QY 22 ProAlaSerProLeuAspSerThrPhe 30  
Db 216 CCTCCAGCCCATGGACAGACCTTC 242

RESULT 2  
AW606421  
LOCUS RC3-HT0371-290100-013-e06 HT0371 Homo sapiens cDNA, mRNA linear EST 23-MAR-2000  
DEFINITION AW606421  
ACCESSION AW606421  
VERSION AW606421.1 GI:7311162  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 422)  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE HCGP <http://www.ludwig.org.br/ORESTES>.  
The FAPESP/LICR Human Cancer Genome Project

# JOURNAL COMMENT

Unpublished  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3&t2=RC3-HT0371-290100-013-e06&t3=2000-01-29&t4=1>)  
Seq primer: puc 18 forward  
High quality sequence start: 8  
High quality sequence stop: 422.  
Location/Qualifiers  
1. .422  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="HT0371"  
/note="Organ: head,neck; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

# FEATURES

source  
1. .422  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="HT0371"  
/note="Organ: head,neck; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 77 a 126 c 135 g 84 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.25 Length: 422  
Score: 85.00 Matches: 23  
Percent Similarity: 46.00% Conservative: 0  
Best Local Similarity: 46.00% Mismatches: 3  
Query Match: 53.46% Indels: 24  
DB: 9 Gaps: 2

JUNC\_SEQ3\_SEQ4 (1-30) x AW606421 (1-422)

QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15  
Db 256 CCTGCAGAGGTGGGAGGAGAGATGATCCAGTATGCCAGGCCCTCACGGAGGCTGC 315  
QY 16 -----GlnAsnGluAspLeu 20  
Db 316 ATGCTGGGTGGGAGGGGCCACCATCTCTCTCTCTCCACAGATGAGGACTTG 375

QY 21 GlyProAlaSerProLeuAspSerThrPhe 30  
Db 376 GGCCCGAGGCGAGTCCCTTGGACAGCACCTTC 405

RESULT 3  
BE157169/c  
LOCUS RC3-HT0371-250200-016-all HT0371 Homo sapiens cDNA, mRNA linear EST 21-JUN-2000  
DEFINITION BE157169  
ACCESSION BE157169  
VERSION BE157169.1 GI:8619890  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 439)  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Matsumura, A., Baia, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

```

"
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20202663
PUBMED     10737800
COMMENT    Contact: Simpson A.J.G.
           Laboratory of Cancer Genetics
           Ludwig Institute for Cancer Research
           Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
           Brazil
           Tel: +55-11-2704922
           Fax: +55-11-2707001
           Email: asimpson@ludwig.org.br
           This sequence was derived from the FAPESP/LICR Human Cancer Genome
           Project. This entry can be seen in the following URL
           (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-RC3-HT0371-250
           200-016-all&t3=2000-02-25&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 439.
Location/Qualifiers
FEATURES   1..439
            source
            1..439
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /dev_stage="Adult"
            /clone_lib="HT0371"
            /note="Organ: head/neck; Vector: puc18; Site_1: SmaI;
            Site_2: SmaI; A mini-library was made by cloning products
            derived from ORESTES PCR (U.S. Letters Patent application
            No. 196,716 - Ludwig Institute for Cancer Research)
            profiles into the pUC 18 vector. Reverse transcription of
            tissue mRNA and cDNA amplification were performed under
            low stringency conditions."
BASE COUNT 89 a 139 c 132 g 79 t
ORIGIN

Alignment Scores:
Pred. No.:      0.262      Length:      439
Score:          85.00      Matches:      23
Percent Similarity: 46.00%      Conservative: 0
Best Local Similarity: 46.00%      Mismatches: 3
Query Match:    53.46%      Indels:      24
DB:             10        Gaps:        2

JUNC_SEQ3_SEQ4 (1-30) x BE157169 (1-439)

Qy      5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
      |||||||
Db      167 CCTGCAGAGGTGGGAGGAGAGATGATGCCAGTATGCCAGGCCCTCACGGAAGGCTGC 108
Qy      16 -----GlnAsnGluAspLeu 20
      |||||||
Db      107 ATGCTGGGCTGGGAGGGGCCACCATCTGCTCTCTCTCTCCACAGATGAGGACTTG 48
Qy      21 GlyProAlaSerProLeuAspSerThrPhe 30
      |||||||
Db      47 GCCCCAGCCATCTCTGGACAGCACCTTC 18

RESULT 4
LOCUS    CD558559
DEFINITION AGNCOURT_14477527 NIH_MGC_181 Homo sapiens cDNA clone
ACCESSION IMAGE:30396701 5', mRNA sequence.
VERSION    CD558559
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 944)
NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)

```

## JOURNAL COMMENT

Unpublished  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgaps-r@mail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein  
cDNA Library Preparation: Invitrogen Corp  
DNA Sequencing by: Agencourt Bioscience Corporation (LLNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: NDAM478 row: n column: 06  
High quality sequence start: 19  
High quality sequence stop: 537.

## FEATURES source

```

1..944
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30396701"
/tissue_type="White Matter"
/dev_stage="Unknown"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_181"
/note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
(destroyed); Library is oligo-dr primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH_MGC Library."
BASE COUNT 150 a 269 c 366 g 158 t 1 others
ORIGIN

```

## Alignment Scores:

```

Pred. No.:      0.682      Length:      944
Score:          85.00      Matches:      23
Percent Similarity: 46.00%      Conservative: 0
Best Local Similarity: 46.00%      Mismatches: 3
Query Match:    53.46%      Indels:      24
DB:             14        Gaps:        2

```

JUNC\_SEQ3\_SEQ4 (1-30) x CD558559 (1-944)

```

Qy      5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
      |||||||
Db      193 CCTGCAGAGGTGGGAGGAGAGATGATGCCAGTATGCCAGGCCCTCACGGAAGGCTGC 252
Qy      16 -----GlnAsnGluAspLeu 20
      |||||||
Db      253 ATGCTGGGCTGGGAGGGGCCACCATCTGCTCTCTCTCTCCACAGATGAGGACTTG 312
Qy      21 GlyProAlaSerProLeuAspSerThrPhe 30
      |||||||
Db      313 GCCCCAGCCATCTCTGGACAGCACCTTC 342

```

## RESULT 5

```

LOCUS    BE157103
DEFINITION RC3-HT0371-180200-015-h07 HT0371 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE157103
VERSION    BE157103.1 GI:8619824
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens

```

BE157103 458 bp mRNA linear EST 21-JUN-2000

RC3-HT0371-180200-015-h07 HT0371 Homo sapiens cDNA, mRNA sequence.

BE157103

BE157103.1 GI:8619824

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 458)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordino, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

**TITLE** Simpson,A.J.  
**JOURNAL** Shotgun sequencing of the human transcriptome with ORF expressed  
**MEDLINE** sequence tags  
**PUBMED** Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
**COMMENT** 20202663  
 10737800  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=&t2=RC3-HT0371-180  
 200-015-h07&t3=2000-02-18&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 71  
 High quality sequence stop: 458.

# FEATURES

source  
 1. .458  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HT0371"  
 /note="Organ: head\_neck; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."  
 BASE COUNT 87 a 136 c 138 g 97 t

Alignment Scores:  
 Pred. No.: 0.894 Length: 458  
 Score: 81.00 Matches: 16  
 Percent Similarity: 88.89% Conservative: 0  
 Best Local Similarity: 88.89% Mismatches: 2  
 Query Match: 50.94% Indels: 0  
 DB: 10 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x BE157103 (1-458)

QY 13 LeuThrSerGlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPhe 30  
 ||| |||||  
 Db 388 CTTCCCTCCACAGAATGAGGACTTGGGCCCGCCAGTCCTTGGACAGCACCTTC 441

**RESULT 6**  
**BF869854/c**  
**LOCUS** BF869854 378 bp mRNA linear EST 17-JAN-2001  
**DEFINITION** IL3-ET0114-251000-317-B07 ET0114 Homo sapiens cDNA, mRNA sequence.  
**ACCESSION** BF869854  
**VERSION** BF869854.1 GI:12259984  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**REFERENCE**  
**AUTHORS** 1 (bases 1 to 378)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.

**TITLE** Shotgun sequencing of the human transcriptome with ORF expressed  
**JOURNAL** sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663  
 10737800  
**COMMENT**

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=IL3&t2=IL3-ET0114-  
 251000-317-B07&t3=2000-10-25&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 177.

# FEATURES

source  
 1. .378  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="ET0114"  
 /note="Organ: lung\_tumor; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."  
 BASE COUNT 76 a 119 c 110 g 73 t

Alignment Scores:  
 Pred. No.: 0.944 Length: 378  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.31% Indels: 0  
 DB: 10 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x BF869854 (1-378)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
 |||||  
 Db 237 GACAAGGGTGGCCCGCCGAGCAGAGCCAGCCCTTGGAGTCC 193

**RESULT 7**  
**HSM069572**  
**ID** HSM069572 standard; RNA; EST; 411 BP.  
**XX**  
**AC** BX479682;  
**XX**  
**SV** BX479682.1  
**XX**  
**DT** 09-MAY-2003 (Rel. 75, Created)  
**DT** 09-MAY-2003 (Rel. 75, Last updated, Version 1)  
**XX**  
**DE** Homo sapiens mRNA; EST DKFZp686K21213\_r1 (from clone DKFZp686K21213)  
**XX**  
**KW** EST; expressed sequence tag.

**OS** Homo sapiens (human)  
**OC** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;  
**OC** Eutheria; Primates; Catarrhini; Hominidae; Homo.

[1]

RP 1-411  
 RA Bloeker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,  
 Han M., Wiemann S.;  
 RT ;  
 RL Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.  
 RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY

XX This is the 5' sequence of the clone insert  
 CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 CC Research Center (DKFZ): Email s.wiemann@dkfz-heidelberg.de;  
 CC sequenced by GBF (National Research Centre for Biotechnology  
 CC Ltd., Braunschweig/Germany) within the cDNA sequencing  
 CC consortium of the German Genome Project.  
 CC No sl sequence available.  
 CC This clone (DKFZp686K21213) is available at the RZPD in Berlin.  
 CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6,  
 CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

XX Key Location/Qualifiers  
 FH source  
 FT 1. 411  
 FT /db\_xref="taxon:9606"  
 FT /mol\_type="mRNA"  
 FT /organism="Homo sapiens"  
 FT /clone\_lib="DKFZp686K21213"  
 FT DH10B; sites SfiIA + SfiIB  
 FT /dev\_stage="adult"  
 FT /tissue\_type="cDNA-collection"

SQ Sequence 411 BP; 70 A; 126 C; 125 G; 90 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.05 Length: 411  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.31% Indels: 0  
 DB: 2 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x HSM069572 (1-411)

Qy 1 AsPLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

Db 292 GACAAGGCTGCCCGCGAGAGAGAGCCAGCCCTCTGACGTC 336

RESULT 8  
 BF762317/c 433 bp mRNA linear EST 12-JAN-2001  
 LOCUS IL2-CS0049-181000-202-A06 CS0049 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION BF762317  
 ACCESSION BF762317  
 VERSION BF762317.1 GI:12110217  
 KEYWORDS EST  
 SOURCE Homo sapiens (human)

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 433)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=IL2&t2=IL2-CS0049-  
 181000-202-A06&t3=2000-10-18&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 387.

#### FEATURES

source

Location/Qualifiers

1..433

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="CS0049"

/note="Organ: colon\_est; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

BASE COUNT 86 a 149 c 108 g 90 t

ORIGIN

Alignment Scores:

Pred. No.: 1.12 Length: 433  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.31% Indels: 0  
 DB: 10 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x BF762317 (1-433)

Qy 1 AsPLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

Db 367 GACAAGGCTGCCCGCGAGAGAGCCAGCCCTCTGACGTC 323

RESULT 9

BG988031

LOCUS MR2-HT1160-110101-005-a06 HT1160 Homo sapiens cDNA, mRNA sequence.

DEFINITION BG988031

ACCESSION BG988031.1 GI:14392101

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 450)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR2&t2=MR2-HT1160-

110101-005-a06&t3=2001-01-11&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 39

High quality sequence stop: 325.

FEATURES  
source

Location/Qualifiers  
1. .450  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="HT1160"  
/note="Organ: head\_neck; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
84 a 127 c 150 g 89 t

BASE COUNT  
ORIGIN

84 a 127 c 150 g 89 t

## Alignment Scores:

Pred. No.: 1.17 Length: 450  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x BG988031 (1-450)

QY 1 AsPLySGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

Db 259 GACAAGGGCTGCCCGCAGAGAGAGCCGCTCTGACGTCC 303

## RESULT 10

## BM720186

## LOCUS

BM720186 466 bp mRNA linear EST 01-MAR-2002  
UI-E-E00-ahz-c-08-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone

UI-E-E00-ahz-c-08-0-UI 5', mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 466)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

8889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse

Location/Qualifiers

1. .466

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="UI-E-E00-ahz-c-08-0-UI"

/tissue\_type="fetal eye"

/dev\_stage="fetal"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone\_lib="UI-E-E00"

/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site\_1: EcoR I; Site\_2: Not I;

UI-E-E00 is a cDNA library containing the following

tissue(s): fetal eye. The library was constructed

according to Bonaldo, Lennon and Soares, Genome Research,

6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dT primer containing a Not I site. Double

stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pT7T3-Pac

vector. The oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The

sequence tag for this library is CGGTATACC. This library

was created for the program, Gene Discovery in the Visual

System, supported by National Eye Institute (NEI)."

System, supported by National Eye Institute (NEI)."

BASE COUNT 105 a 114 c 151 g 93 t

ORIGIN

## Alignment Scores:

Pred. No.: 1.23 Length: 466  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x BM720186 (1-466)

QY 1 AsPLySGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

Db 18 GACAAGGGCTGCCCGCAGAGAGAGCCGCTCTGACGTCC 62

## RESULT 11

## BI194790

## LOCUS

## DEFINITION

BI194790 546 bp mRNA linear EST 10-JUL-2001

602948150F1 NIH\_MGC\_42 Homo sapiens cDNA clone IMAGE:5091396 5',

mRNA sequence.

ACCESSION BI194790.1 GI:14649810

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 546)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@email.nih.gov

Tissue Procurement: AFCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCMI855 row: c column: 13

High quality sequence start: 10

High quality sequence stop: 495.

Location/Qualifiers

1. .546

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5091396"

/tissue\_type="epithelioid carcinoma cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_42"

/note="Organ: pancreas; Vector: pOTB7; Site\_1: XhoI;

Site\_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the



following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library. |"

BASE COUNT 110 a 147 c 183 g 106 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.49 Length: 546  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x BI194790 (1-546)

Qy 1 AsPLysGLyCysProAlaGLuInArgAlaSerProLeuThrSer 15  
|||||  
Db 223 GACAAGGCTGCCCGCGAGCAGAGCCGCTCTGAGCTCC 267

RESULT 12

AA443351  
LOCUS zw85b05.r1 Soares\_total\_fetus\_Nb2HF8.9w Homo sapiens cDNA clone  
DEFINITION IMAGE:783729 5' similar to gb:M11730 ERBB-2 RECEPTOR  
PROTEIN-TYROSINE KINASE PRECURSOR (HUMAN); mRNA sequence.

ACCESSION AA443351

VERSION AA443351.1 GI:2156026

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,  
Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie,  
T., Waterston, R., and Wilson, R.

TITLE Washu-Merck EST Project 1997

JOURNAL Unpublished

COMMENT

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 360.

Location/Qualifiers

FEATURES

source

1..547  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:5981347"  
/db\_xref="taxon:9606"  
/clone="IMAGE:783729"  
/dev\_stage="8-9 weeks"  
/lab\_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGCGCTTAATTTTTTTTTTTT 3'].  
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 115 a 151 c 172 g 108 t  
ORIGIN

Alignment Scores:

Pred. No.: 1.5 Length: 547  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 9 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x AA443351 (1-547)

Qy 1 AsPLysGLyCysProAlaGLuInArgAlaSerProLeuThrSer 15  
|||||  
Db 174 GACAAGGCTGCCCGCGAGCAGAGCCGCTCTGAGCTCC 218

RESULT 13

BM720098

LOCUS

DEFINITION UI-E-E00-ahz-c-05-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone  
UI-E-E00-ahz-c-05-0-UI 5', mRNA sequence.

ACCESSION BM720098

VERSION BM720098.1 GI:19039119

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

Location/Qualifiers

FEATURES

source

1..561  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-E00-ahz-c-05-0-UI"  
/tissue\_type="fetal eye"  
/dev\_stage="fetal"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-E00"

/note="Organ: eye; Vector: pT7T3-pac (Pharmacia) with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I; UI-E-E00 is a cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGCGTATACC. This library

was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

```

BASE COUNT      97 a 171 c 179 g 113 t 1 others
ORIGIN

Alignment Scores:
Pred. No.:      1.54      Length:      561
Score:          80.00     Matches:      15
Percent Similarity: 100.00%  Conservatives: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    50.31%   Indels:      0
DB:            12      Gaps:      0

JUNC_SEQ3_SEQ4 (1-30) x BM720098 (1-561)

QY      1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
      |||||
Db      400 GACAAGGGCTGCCCGCGAGCAGAGACCGCCCTCTGACGTC 444

RESULT 14
BM083444
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BM083444
imageQC_6-2000/sjp459bdf42.x2 NIH_MGC_39 Homo sapiens cDNA clone
IMAGE:3610674 5', mRNA sequence.
BM083444
BM083444.1 GI:16951075
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Kale,P.I., Harsch,T.J., Folta,P.A., Nelson,D.O., Sanders,C.G. and
Prange,C.K.
The I.M.A.G.E. Consortium quality control effort: clone
Resequencing for verification
Unpublished
Other_ESTs: BB615590
Contact: Prange CK
The I.M.A.G.E. Consortium
Lawrence Livermore National Laboratory
Livermore, CA, USA
Email: help@image.llnl.gov
This read has been verified (found to hit its original self in the
correct orientation), as part of the I.M.A.G.E. Consortium quality
control effort. High quality sequence is defined as having 100 or
more base pairs with a phred quality value of 20 or greater, where
a sliding window of 4 base pairs with a phred quality value of 15
or greater marks the beginning and end of the sequence. For
information on obtaining this clone, please contact
info@image.llnl.gov.
Plate: LLCM267 row: b column: 19
Seq primer: -21m13
High quality sequence stop: 607.
Location/Qualifiers
1..607
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3610674"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_39"
/site="Organ: pancreas; Vector: pOTB7; Site.1: XhoI;
Site.2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies)."

BASE COUNT      118 a 171 c 195 g 121 t 2 others
ORIGIN

```

## Alignment Scores:

```

Pred. No.:      1.7      Length:      607
Score:          80.00     Matches:      15
Percent Similarity: 100.00%  Conservatives: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    50.31%   Indels:      0
DB:            12      Gaps:      0

```

JUNC\_SEQ3\_SEQ4 (1-30) x BM083444 (1-607)

```

QY      1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
      |||||
Db      292 GACAAGGGCTGCCCGCGAGCAGAGACCGCCCTCTGACGTC 336

```

## RESULT 15

AW410534

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AW410534

fh06h06.xl NIH\_MGC\_17 Homo sapiens cDNA clone IMAGE:2961635 5',

mRNA sequence.

AW410534

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/dbrrp/image/image.html

Plate: LLCM56 row: 0 column: 12

Seq primer: -21m13 forward primer (ABI).

Location/Qualifiers

1..613

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:2961635"

/tissue\_type="rhabdomyosarcoma"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_17"

/note="Organ: muscle; Vector: pOTB7; Site.1: EcoRI;

Site.2: XhoI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Size-selected >500bp

for average insert size 1.8kb. Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 109 a 205 c 185 g 113 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 1.73 Length: 613

Score: 80.00 Matches: 15

Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 50.31% Indels: 0

DB: 9 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x AW410534 (1-613)

Qy 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
Db 557 GACAAGGGCTGCCCCCGCGAGCAGAGAGCCAGCCCTCTGACGTCC 601

Search completed: October 16, 2003, 17:04:02  
Job time : 237.794 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 16, 2003, 11:08:07 : Search time 31.0111 Seconds  
(without alignments)  
2540.503 Million cell updates/sec

Title: JUNC\_SEQ3\_SEQ4

Perfect score: 159

Sequence: 1 DKGPCAPQASPLTSQNEGLGASPLDSTF 30

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-O/cgn2\_1/USPTO\_spool\_p/HOLLERAN80/runat\_15102003\_131915\_20662/app\_query.fasta\_1.4685  
-DB=Published\_Applications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=HOLLERAN80 -CGN\_1\_1\_397=runat\_15102003\_131915\_20662  
-NCPU=6 -ICPU=3 -NO\_WMAP -LANGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Published Applications NA:\*
- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
  - 2: /cgn2\_6/ptodata/2/pubpna/PCCT\_NEW\_PUB.seq:\*
  - 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
  - 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
  - 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
  - 6: /cgn2\_6/ptodata/2/pubpna/PCCTUS\_PUBCOMB.seq:\*
  - 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
  - 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
  - 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
  - 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
  - 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
  - 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
  - 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
  - 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
  - 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
  - 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
  - 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	50.3	201	13	US-10-109-213-3
Sequence 3, Appli					

Alignment Scores:	0.000724	Length:	201
Pred. No.:	80.00	Matches:	15
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%		

ALIGNMENTS

RESULT 1

US-10-109-213-3  
; Sequence 3, Application US/10109213  
; Publication No. US20020186670A1  
; GENERAL INFORMATION:  
; APPLICANT: Ecker, David J.  
; TITLE OF INVENTION: Identification Of Disease Predictive Nucleic Acids  
; FILE REFERENCE: IBIS0009  
; CURRENT APPLICATION NUMBER: US/10/109,213  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: US/09/200,355  
; PRIOR FILING DATE: 1998-11-25  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 201  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-10-109-213-3

2	80	50.3	3765	12	US-10-207-498-5	Sequence 5, Appli
3	80	50.3	3768	9	US-09-811-123-8	Sequence 8, Appli
4	80	50.3	3768	9	US-09-811-115-2	Sequence 2, Appli
5	80	50.3	3768	10	US-09-854-356-9	Sequence 9, Appli
6	80	50.3	3768	10	US-09-930-125-1	Sequence 1, Appli
7	80	50.3	3768	12	US-10-313-644-1	Sequence 1, Appli
8	80	50.3	4473	11	US-09-441-411-5	Sequence 5, Appli
9	80	50.3	4473	12	US-10-101-510-81	Sequence 81, Appli
10	80	50.3	4473	14	US-10-146-473-32	Sequence 32, Appli
11	80	50.3	4473	14	US-10-207-655-44	Sequence 44, Appli
12	80	50.3	4530	10	US-09-877-177-11	Sequence 11, Appli
13	80	50.3	4530	12	US-10-087-926A-119	Sequence 119, App
14	80	50.3	4530	12	US-10-101-510-124	Sequence 124, App
15	80	50.3	4530	12	US-10-338-730-1	Sequence 1, Appli
16	80	50.3	4530	14	US-10-177-293-125	Sequence 125, App
17	80	50.3	4543	10	US-09-769-508-1	Sequence 1, Appli
18	80	50.3	4606	12	US-09-971-392-70	Sequence 70, Appli
19	80	50.3	4642	14	US-10-198-846-10896	Sequence 10896, A
20	80	50.3	9274	9	US-09-811-123-7	Sequence 7, Appli
21	80	50.3	9274	9	US-09-811-115-1	Sequence 1, Appli
22	79	49.7	1115	14	US-10-102-806-165	Sequence 165, App
23	79	49.7	1713	12	US-10-378-393-14	Sequence 14, Appli
24	79	49.7	1755	10	US-09-930-125-6	Sequence 6, Appli
25	79	49.7	1767	10	US-09-930-125-4	Sequence 4, Appli
26	79	49.7	1773	10	US-09-930-125-7	Sequence 7, Appli
27	79	49.7	1806	10	US-09-930-125-5	Sequence 5, Appli
28	79	49.7	2411	12	US-10-378-393-10	Sequence 10, Appli
29	74	46.5	3771	10	US-09-854-356-11	Sequence 11, Appli
30	74	46.5	3955	10	US-09-870-759-117	Sequence 117, App
31	74	46.5	3955	10	US-09-854-356-10	Sequence 10, Appli
32	74	46.5	3955	12	US-09-751-708A-117	Sequence 117, App
33	62	39.0	14427	14	US-10-156-761-1540	Sequence 1540, Ap
34	62	39.0	23432	9	US-09-764-869-1332	Sequence 1332, Ap
35	62	39.0	23432	14	US-10-091-504-1332	Sequence 1332, Ap
36	62	39.0	9025608	14	US-10-156-761-1	Sequence 1, Appli
37	59	37.1	111822	13	US-10-094-989-3	Sequence 3, Appli
38	58.5	36.8	633	11	US-09-788-188-12	Sequence 12, Appli
39	58.5	36.8	634	8	US-08-450-842-1	Sequence 1, Appli
40	58.5	36.8	1404	14	US-10-155-785-13	Sequence 13, Appli
41	57.5	36.2	363	11	US-09-918-995-28706	Sequence 28706, A
42	57	35.8	297	10	US-09-783-590-11498	Sequence 11498, A
43	57	35.8	1938	12	US-10-422-264-25	Sequence 25, Appli
44	57	35.8	2316	12	US-10-422-264-27	Sequence 27, Appli
45	57	35.8	2604	12	US-10-422-264-23	Sequence 23, Appli

Query Match: 50.31% Indels: 0  
DB: 13 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x US-10-109-213-3 (1-201)

QY 1 AsPLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
DB 116 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCUCGACGTC 160

## RESULT 2

US-10-207-498-5  
; Sequence 5, Application US/10207498  
; Publication No. US20030143568A1  
; GENERAL INFORMATION:  
; APPLICANT: Elizabeth Singer  
; APPLICANT: Ralf Landgraf  
; APPLICANT: Dennis J. Slamon  
; APPLICANT: David Eisenberg  
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING  
; FILE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HEREGULIN AND HER3  
; FILE REFERENCE: 30448.103-US-01  
; CURRENT APPLICATION NUMBER: US/10/207,498  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 60/308,431  
; PRIOR FILING DATE: 2001-07-27  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 3765  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(3765)  
US-10-207-498-5

Alignment Scores:  
Pred. No.: 0.0177 Length: 3765  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x US-10-207-498-5 (1-3765)

QY 1 AsPLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
DB 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTC 1959

## RESULT 3

US-09-811-123-8  
; Sequence 8, Application US/09811123  
; Patent No. US20020001587A1  
; GENERAL INFORMATION:  
; APPLICANT: Sharon Erickson  
; APPLICANT: Ralph Schwall  
; APPLICANT: Mark Sliwowski  
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB  
; FILE REFERENCE: GENENT.073A2  
; CURRENT APPLICATION NUMBER: US/09/811,123  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/238,327  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 09/602,530  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 3768  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-811-123-8

## Alignment Scores:

Pred. No.: 0.0177 Length: 3768  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 9 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x US-09-811-123-8 (1-3768)

QY 1 AsPLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
DB 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTC 1959

## RESULT 4

US-09-811-115-2  
; Sequence 2, Application US/09811115  
; Patent No. US20020035736A1  
; GENERAL INFORMATION:  
; APPLICANT: Erickson, Sharon  
; APPLICANT: Schwall, Ralph  
; APPLICANT: King, Kathleen  
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL  
; FILE REFERENCE: GENENT.034A  
; CURRENT APPLICATION NUMBER: US/09/811,115  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/189,844  
; PRIOR FILING DATE: 2000-03-16  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 3768  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-811-115-2

Alignment Scores:  
Pred. No.: 0.0177 Length: 3768  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 9 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x US-09-811-115-2 (1-3768)

QY 1 AsPLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
DB 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTC 1959

## RESULT 5

US-09-854-356-9  
; Sequence 9, Application US/09854356  
; Patent No. US20020177567A1  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Gheysen, Dirk  
; APPLICANT: Corixa Corporation  
; APPLICANT: SmithKline Beecham Biologicals S. A.  
; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
; FILE REFERENCE: 014058-009810PC  
; CURRENT APPLICATION NUMBER: US/09/854,356  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: US 09/493,480  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/117,976  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 3768

Alignment Scores:	
Pred. No.:	0.0214
Score:	0.000
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	50.31%
DB:	11
Length:	4473
Matches:	15
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 2089 GACAAGGGCTGCCCCGCGAGCAGAGAGCCAGCCCTCTGACGTCC 2133

## RESULT 9

US-10-101-510-81

; Sequence 81, Application US/10101510  
; Publication No. US20030148295A1  
; GENERAL INFORMATION:  
; APPLICANT: WAN, JACKSON  
; APPLICANT: WANG, YIXIN  
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE  
; FILE REFERENCE: 15117.0012  
; CURRENT APPLICATION NUMBER: US/10/101,510  
; CURRENT FILING DATE: 2002-03-20  
; PRIOR APPLICATION NUMBER: 60/276,947  
; PRIOR FILING DATE: 2001-03-20  
; NUMBER OF SEQ ID NOS: 805  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 81  
; LENGTH: 4473  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-101-510-81

Alignment Scores:  
Pred. No.: 0.0214 Length: 4473  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 50.31%  
Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x US-10-101-510-81 (1-4473)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 2089 GACAAGGGCTGCCCCGCGAGCAGAGAGCCAGCCCTCTGACGTCC 2133

## RESULT 10

US-10-146-473-32

; Sequence 32, Application US/10146473  
; Publication No. US2003010888A1  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Matthew  
; APPLICANT: Gout, Ivan  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Gure, Ali  
; APPLICANT: Chen, Yao-Tseng  
; APPLICANT: Old, Lloyd  
; TITLE OF INVENTION: Breast Cancer Antigens  
; FILE REFERENCE: L00461/70130(JRV)  
; CURRENT APPLICATION NUMBER: US/10/146,473  
; CURRENT FILING DATE: 2002-05-15  
; PRIOR APPLICATION NUMBER: US 60/291,150  
; PRIOR FILING DATE: 2001-05-15  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 32  
; LENGTH: 4473  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-146-473-32

Alignment Scores:  
Pred. No.: 0.0214 Length: 4473  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 50.31%  
Indels: 0  
DB: 14 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x US-10-146-473-32 (1-4473)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 2089 GACAAGGGCTGCCCCGCGAGCAGAGAGCCAGCCCTCTGACGTCC 2133

## RESULT 11

US-10-207-655-44

; Sequence 44, Application US/10207655  
; Publication No. US20030118592A1  
; GENERAL INFORMATION:  
; APPLICANT: Ledbetter, Jeffrey A.  
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
; FILE REFERENCE: 390069, 401C1  
; CURRENT APPLICATION NUMBER: US/10/207,655  
; CURRENT FILING DATE: 2002-07-25  
; NUMBER OF SEQ ID NOS: 426  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 44  
; LENGTH: 4473  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-207-655-44

Alignment Scores:  
Pred. No.: 0.0214 Length: 4473  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 50.31%  
Indels: 0  
DB: 14 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x US-10-207-655-44 (1-4473)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 2089 GACAAGGGCTGCCCCGCGAGCAGAGAGCCAGCCCTCTGACGTCC 2133

## RESULT 12

US-09-877-177-11

; Sequence 11, Application US/09877177  
; Publication No. US20020192652A1  
; GENERAL INFORMATION:  
; APPLICANT: Peter V. Danenberg et al.  
; TITLE OF INVENTION: Method of determining Epidermal Growth  
; FILE REFERENCE: 11220/120  
; CURRENT APPLICATION NUMBER: US/09/877,177  
; CURRENT FILING DATE: 2001-06-11  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 4530  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-877-177-11

Alignment Scores:  
Pred. No.: 0.0217 Length: 4530  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 50.31%  
Indels: 0  
DB: 10 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x US-09-877-177-11 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 2065 GACAAGGGCTGCCCCGCGAGCAGAGAGCCAGCCCTCTGACGTCC 2109

## RESULT 13

US-10-007-926A-119





THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: October 15, 2003, 22:55:19 ; Search time 7.60488 Seconds  
(without alignments)  
1741.185 Million cell updates/sec

Title: JUNC\_SEQ3\_SEQ4  
Perfect score: 159  
Sequence: 1 DKGPCAQRASPLTSQNEGLGPASPLDSTF 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -MODEL=frame\_plus\_p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/HOLLERAN480/runat\_15102003\_131913\_20570/app\_query.fasta\_1.4685  
-DB=Issued\_Patents\_NA -QMT=Fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTEXT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 -CGN\_1\_1\_403=runat\_15102003\_131913\_20570 -NCPU=6 -ICPU=3  
-DEV\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents, NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	84	52.8	2385 2 US-09-146-283-3 Sequence 3, Appli
2	84	52.8	2385 3 US-08-579-823A-3 Sequence 3, Appli
3	84	52.8	2385 3 US-09-344-195-3 Sequence 3, Appli
4	80	50.3	153 3 US-08-776-251-3 Sequence 3, Appli
5	80	50.3	201 4 US-09-200-355-3 Sequence 10, Appli
6	80	50.3	816 3 US-08-776-251-10 Sequence 1, Appli
7	80	50.3	3768 2 US-08-625-101-1 Sequence 1, Appli
8	80	50.3	3768 2 US-08-356-786-1 Sequence 1, Appli
9	80	50.3	4473 2 US-09-048-804-1 Sequence 1, Appli
10	80	50.3	4473 3 US-09-056-105-26 Sequence 26, Appli
11	80	50.3	4530 1 US-08-229-515A-9 Sequence 9, Appli
12	80	50.3	4530 1 US-08-645-865-9 Sequence 9, Appli

13	80	50.3	4530 4 US-09-167-322-4 Sequence 4, Appli
14	80	50.3	4530 4 US-09-527-487-1 Sequence 1, Appli
15	80	50.3	4530 4 US-09-877-177A-11 Sequence 11, Appli
16	74	46.5	3955 1 US-08-229-515A-14 Sequence 14, Appli
17	74	46.5	3955 1 US-08-645-865-14 Sequence 14, Appli
c 18	59	37.1	111282 4 US-09-754-250-3 Sequence 3, Appli
c 19	58.5	36.8	634 1 US-08-451-947-1 Sequence 1, Appli
c 20	58.5	36.8	634 2 US-08-424-826A-1 Sequence 1, Appli
c 21	58.5	36.8	634 3 US-08-928-694-1 Sequence 1, Appli
c 22	58.5	36.8	634 4 US-08-450-842-1 Sequence 1, Appli
c 23	58.5	36.8	634 4 US-08-451-390-1 Sequence 1, Appli
c 24	58.5	36.8	634 5 PCT-US91-06950-1 Sequence 1, Appli
c 25	58.5	36.8	1404 1 US-07-796-106-22 Sequence 22, Appli
c 26	58.5	36.8	4403765 3 US-09-103-840A-2 Sequence 2, Appli
c 27	58.5	36.8	4411529 3 US-09-103-840A-1 Sequence 1, Appli
c 28	57	35.8	1938 4 US-09-547-435-25 Sequence 25, Appli
c 29	57	35.8	2316 4 US-09-547-435-27 Sequence 27, Appli
c 30	57	35.8	2604 4 US-09-547-435-23 Sequence 23, Appli
c 31	57	35.8	3384 4 US-09-547-435-29 Sequence 29, Appli
c 32	55.5	34.9	2691 4 US-09-020-743-1 Sequence 1, Appli
c 33	55	34.6	960 3 US-08-651-136C-1 Sequence 1, Appli
c 34	55	34.6	960 4 US-09-229-911A-1 Sequence 1, Appli
c 35	55	34.6	4092 3 US-09-306-595C-5 Sequence 5, Appli
c 36	55	34.6	4092 4 US-09-925-388-5 Sequence 5, Appli
c 37	55	34.6	5046 4 US-09-548-938A-5 Sequence 5, Appli
c 38	54.5	34.3	824 4 US-09-312-283C-369 Sequence 369, App
c 39	53.5	33.6	685 1 US-08-451-947-7 Sequence 7, Appli
c 40	53.5	33.6	685 2 US-08-424-826A-7 Sequence 7, Appli
c 41	53.5	33.6	685 3 US-08-928-694-7 Sequence 7, Appli
c 42	53.5	33.6	685 4 US-08-450-842-7 Sequence 7, Appli
c 43	53.5	33.6	685 4 US-08-451-390-7 Sequence 7, Appli
c 44	53.5	33.6	685 5 PCT-US91-06950-7 Sequence 7, Appli
c 45	53.5	33.6	1190 1 US-08-451-947-9 Sequence 9, Appli

ALIGNMENTS

RESULT 1  
US-09-146-283-3  
; Sequence 3, Application US/09146283  
; Patent No. 5976346  
; GENERAL INFORMATION:  
; APPLICANT: Laus, Reiner  
; APPLICANT: Ruegg, Curtis L.  
; APPLICANT: Wu, Hongyu  
; TITLE OF INVENTION: Immunostimulatory Compositions  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Ave. Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/146,283  
; FILING DATE: 03-SEPT-1998  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Judge, Linda R.  
; REGISTRATION NUMBER: 42,702  
; REFERENCE/DOCKET NUMBER: 7636-0010.21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-0880  
; TELEFAX: 650-324-0960  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2385 base pairs

```
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-09-146-283-3

Alignment Scores:
Pred. No.: 0.0053 Length: 2385
Score: 84.00 Matches: 21
Percent Similarity: 70.00% Conservative: 0
Best Local Similarity: 70.00% Mismatches: 5
Query Match: 52.83% Indels: 4
DB: 2 Gaps: 2

JUNC_SEQ3_SEQ4 (1-30) x US-09-146-283-3 (1-2385)
Qy 1 AsLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1925 GACAAGGGCTGCCCGCCGAGAGAGAGCCGCTCTGACGCTCCTCGAG----- 1975
Qy 21 GlyProAla---SerProLeuAspSerThr 29
Db 1976 GCACCGCCGCTCGCCAGCCAGCCAGCACA 2005

RESULT 2
US-08-579-823A-3
; Sequence 3, Application US/08579823A
; Patent No. 6080409
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Composition and Method
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94305
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,823A
; FILING DATE: 03-DEC-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-146-283-3

Alignment Scores:
Pred. No.: 0.0053 Length: 2385
Score: 84.00 Matches: 21
Percent Similarity: 70.00% Conservative: 0
Best Local Similarity: 70.00% Mismatches: 5
Query Match: 52.83% Indels: 4
DB: 2 Gaps: 2

JUNC_SEQ3_SEQ4 (1-30) x US-09-146-283-3 (1-2385)
Qy 1 AsLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1925 GACAAGGGCTGCCCGCCGAGAGAGCCGCTCTGACGCTCCTCGAG----- 1975
Qy 21 GlyProAla---SerProLeuAspSerThr 29
Db 1976 GCACCGCCGCTCGCCAGCCAGCCAGCACA 2005

RESULT 3
US-09-344-195-3
; Sequence 3, Application US/09344195
; Patent No. 6210662
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,195
; FILING DATE: 24-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-344-195-3

Alignment Scores:
Pred. No.: 0.0053 Length: 2385
Score: 84.00 Matches: 21
Percent Similarity: 70.00% Conservative: 0
Best Local Similarity: 70.00% Mismatches: 5
Query Match: 52.83% Indels: 4
DB: 2 Gaps: 2

JUNC_SEQ3_SEQ4 (1-30) x US-08-579-823A-3 (1-2385)
Qy 1 AsLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1925 GACAAGGGCTGCCCGCCGAGAGAGCCGCTCTGACGCTCCTCGAG----- 1975
Qy 21 GlyProAla---SerProLeuAspSerThr 29
Db 1976 GCACCGCCGCTCGCCAGCCAGCCAGCACA 2005
```

Pred. No.: 0.0053 Length: 2385  
Score: 84.00 Matches: 21  
Percent Similarity: 70.00% Conservative: 0  
Best Local Similarity: 70.00% Mismatches: 5  
Query Match: 52.83% Indels: 4  
DB: 3 Gaps: 2

JUNC\_SEQ3\_SEQ4 (1-30) x US-09-344-195-3 (1-2385)

Qy 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspleu 20  
Db 1925 GACAAGGCTGCCCGCGAGAGAGAGCCAGCCCTCTGAGCTCCCTCGAG----- 1975  
Qy 21 GlyProAla---SerProLeuAspSerThr 29  
Db 1976 GCACCCGCGCTCGCCCGAGCCCGCCAGCAC 2005

## RESULT 4

US-08-776-251-3  
; Sequence 3, Application US/08776251  
; Patent No. 6025340  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Caroline J  
; APPLICANT: Marais, Richard  
; TITLE OF INVENTION: Surface expression of enzyme in gene directed prodrgug therapy  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye  
; STREET: 1100 No. 6025340th Glebe Road, 8th Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: USA  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,251  
; FILING DATE: 31-JAN-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB95/01782  
; FILING DATE: 27-JUL-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9415167.7  
; FILING DATE: 27-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arthur R. Crawford  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 620-20  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 153 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-776-251-3

Alignment Scores:  
Pred. No.: 0.000654 Length: 153  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 3 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x US-08-776-251-3 (1-153)

Qy 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
Db 10 GACAAGGCTGCCCGCGAGAGAGCCAGCCCTCTGAGCTCC 54

## RESULT 5

US-09-200-355-3  
; Sequence 3, Application US/09200355  
; Patent No. 6451524  
; GENERAL INFORMATION:  
; APPLICANT: Ecker, David J.  
; TITLE OF INVENTION: Identification Of Disease Predictive Nucleic Acids  
; FILE REFERENCE: IBIS0009  
; CURRENT APPLICATION NUMBER: US/09/200,355  
; CURRENT FILING DATE: 1998-11-25  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 3  
; LENGTH: 201  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-200-355-3

Alignment Scores:  
Pred. No.: 0.000922 Length: 201  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.31% Indels: 0  
DB: 4 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x US-09-200-355-3 (1-201)

Qy 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
Db 116 GACAAGGCGCGCCCGAGAGAGCCAGCCCGCUGAGGUCC 160

## RESULT 6

US-08-776-251-10  
; Sequence 10, Application US/08776251  
; Patent No. 6025340  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Caroline J  
; APPLICANT: Marais, Richard  
; TITLE OF INVENTION: Surface expression of enzyme in gene directed prodrgug there  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon & Vanderhye  
; STREET: 1100 No. 6025340th Glebe Road, 8th Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: USA  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,251  
; FILING DATE: 31-JAN-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB95/01782  
; FILING DATE: 27-JUL-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9415167.7  
; FILING DATE: 27-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arthur R. Crawford  
; REGISTRATION NUMBER: 25,327  
; REFERENCE/DOCKET NUMBER: 620-20  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 816 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-776-251-10

Alignment Scores:  
 Pred. No.: 0.00536 Length: 816  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 50.31%  
 DB: 3 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x US-08-776-251-10 (1-816)

QY 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15  
 |||||  
 Db 93 GACAAGGCTGCCCGCCGAGAGAGAGCCGCTCTGACGTCC 137

# RESULT 7

US-08-625-101-1  
 ; Sequence 1, Application US/08625101  
 ; Patent No. 5869445  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cheever, Martin A.  
 ; APPLICANT: Disis, Mary L.  
 ; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE  
 ; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION  
 ; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu  
 ; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SEED AND BERRY LLP  
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue  
 ; CITY: Seattle  
 ; STATE: Washington  
 ; COUNTRY: USA  
 ; ZIP: 98104-7092

# COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/625,101  
 FILING DATE: 01-APR-1996  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sharkey, Richard G.  
 REGISTRATION NUMBER: 32,629  
 REFERENCE/DOCKET NUMBER: 920010.448C7  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3768 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..3765  
 US-08-625-101-1

Alignment Scores:  
 Pred. No.: 0.0366 Length: 3768  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 50.31%  
 DB: 2 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x US-08-625-101-1 (1-3768)

QY 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15  
 |||||

Db 1915 GACAAGGCTGCCCGCCGAGAGAGAGCCGCTCTGACGTCC 1959

# RESULT 8

US-08-356-786-1  
 ; Sequence 1, Application US/08356786  
 ; Patent No. 5877305  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Huston, James S.  
 ; APPLICANT: Oppermann, Hermann  
 ; APPLICANT: Houston, L. L.  
 ; APPLICANT: Ring, David B.  
 ; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
 ; TITLE OF INVENTION: Marker  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
 ; STREET: Exchange Place, 53 State Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/356,786  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/831,967  
 FILING DATE: 06-FEB-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pitcher, Edmund R.  
 REGISTRATION NUMBER: 27,829  
 REFERENCE/DOCKET NUMBER: CRP-053  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 248-7000  
 TELEFAX: (617) 248-7100

# INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 3768 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..3768  
 OTHER INFORMATION: /note= "product = "cerb-b2"  
 US-08-356-786-1

Alignment Scores:  
 Pred. No.: 0.0366 Length: 3768  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 50.31%  
 DB: 2 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x US-08-356-786-1 (1-3768)

QY 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15  
 |||||

Db 1915 GACAAGGCTGCCCGCCGAGAGAGAGCCGCTCTGACGTCC 1959

# RESULT 9

US-09-048-804-1  
 ; Sequence 1, Application US/09048804  
 ; Patent No. 5968748  
 ; GENERAL INFORMATION:  
 ; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters

Score:	80.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	50.31%	Indels:	0
DB:	3	Gaps:	0
JUNC_SEQ3_SEQ4 (1-30) x US-09-056-105-26 (1-4473)			
QY	1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15		
Db	2089 GACAAGGGCTGCCCGCCGACAGACAGACCCCTCTGAGCGTCC 2133		
RESULT 11			
US-08-229-515A-9			
; Sequence 9, Application US/08229515A			
; Patent No. 551885			
; GENERAL INFORMATION:			
; APPLICANT: RAZIUDDIN			
; APPLICANT: SARKAR, FAZLUL H			
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN			
; TITLE OF INVENTION: NEOPLASTIC DISEASE			
; NUMBER OF SEQUENCES: 19			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: NEEDLE & ROSENBERG PC			
; STREET: 127 Peachtree Street, Suite 1200			
; CITY: Atlanta			
; STATE: Georgia			
; COUNTRY: usa			
; ZIP: 30303			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/229,515A			
; FILING DATE: 19 APR 1994			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: PERRYMAN, DAVID G			
; REGISTRATION NUMBER: 33,438			
; REFERENCE/DOCKET NUMBER: 1414.608			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 404-688-0770			
; TELEFAX: 404-688-9880			
; INFORMATION FOR SEQ ID NO: 9:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 4530 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
US-08-229-515A-9			
Alignment Scores:			
Pred. No.:	0.0462	Length:	4530
Score:	80.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	50.31%	Indels:	0
DB:	1	Gaps:	0
JUNC_SEQ3_SEQ4 (1-30) x US-08-229-515A-9 (1-4530)			
QY	1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15		
Db	2065 GACAAGGGCTGCCCGCCGACAGACAGACCCCTCTGAGCGTCC 2109		
RESULT 12			
US-08-645-865-9			
; Sequence 9, Application US/08645865			
; Patent No. 5654406			
; GENERAL INFORMATION:			
; APPLICANT: RAZIUDDIN			

```
;
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-645-865-9

Alignment Scores:
Pred. No.: 0.0462 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-08-645-865-9 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db 2065 GACAGGGGCTGCCCGCCGAGCAGAGCCGCTCTGACGTCC 2109

RESULT 13
US-09-167-322-4
; Sequence 4, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
```

```
;
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-167-322-4

Alignment Scores:
Pred. No.: 0.0462 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-09-167-322-4 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db 2065 GACAGGGGCTGCCCGCCGAGCAGAGCCGCTCTGACGTCC 2109

RESULT 14
US-09-527-487-1
; Sequence 1, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(3915)
US-09-527-487-1

Alignment Scores:
Pred. No.: 0.0462 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x US-09-527-487-1 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db 2065 GACAGGGGCTGCCCGCCGAGCAGAGCCGCTCTGACGTCC 2109

RESULT 15
US-09-877-177A-11
; Sequence 11, Application US/09877177A
```





THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

and is derived by analysis of the total score distribution.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 28.4941 Seconds  
(without alignments)  
2842.104 Million cell updates/sec

Title: JUNC\_SEQ3\_SEQ4  
Perfect score: 159  
Sequence: 1 DKGCPAEQRASPLTSQNEGLGPASPLDSTF 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q/cygn2\_1/USPTO\_spool\_P/HOLLERAN480/runat\_15102003\_131912\_20526/app\_query.fasta\_1.4685  
-DB=N\_Geneseq\_19Jun03 -Qfmt=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosom62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480\_@CEN\_1\_1\_1758\_@runat\_15102003\_131912\_20526 -NCPU=6 -ICPU=3  
-NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_19Jun03.\*

1:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	144	90.6	2763	24	ABA92252 Mouse Her-2/neu ex
2	144	90.6	2781	24	ABA92253 Her-2/neu extracel
3	84	52.8	2385	18	AA772725 Her2-GM-CSF immuno
4	80	50.3	69	21	AA414815 DNA encoding a wil
5	80	50.3	201	21	AA62125 Her2 normal form t
6	80	50.3	2125	19	AAV21727 Humanised vector p
7	80	50.3	2871	21	AA250586 DC8scFv-erbB2EC fu
8	80	50.3	3600	21	AA897136 Human HER-2/neu co
9	80	50.3	3678	24	ABK86207 DNA encoding huma
10	80	50.3	3768	17	AA40739 HER-2/neu oncogene
11	80	50.3	3768	20	AA401912 Human HER-2/neu on
12	80	50.3	3768	21	AA09455 Human heregulin 2
13	80	50.3	3768	22	AAH23392 Human HER-2/neu pr
14	80	50.3	3768	24	AB235744 Human ERBB2 polyinu
15	80	50.3	3768	24	ABX09987 Human ERBB2 DNA fr
16	80	50.3	3768	24	AA43935 Human HER-2 CDNA.
17	80	50.3	3768	24	AA43986 Human ERBB2 antigen
18	80	50.3	3768	24	ABV78168 Human ERBB2 DNA SE
19	80	50.3	3768	24	AA32743 Human Her-2/neu pr
20	80	50.3	3768	24	AA32250 Human Her-2/neu CD
21	80	50.3	3768	24	ABK10730 Human Her-2/neu DN
22	80	50.3	3768	24	ABU91709 Human polynucleoti
23	80	50.3	3768	24	ABK14058 Human HER2 (ErBB2)
24	80	50.3	4299	14	AA046083 Sequence encoding the
25	80	50.3	4472	21	AA414812 CDNA encoding the
26	80	50.3	4473	19	ABQ76220 Human tumour anti
27	80	50.3	4473	20	AB231071 HER-2 nucleic acid
28	80	50.3	4473	24	AB234969 Human gene express
29	80	50.3	4473	24	AA38904 Human Her-2 DNA.
30	80	50.3	4530	16	AA01585 Her-2/neu (ERBB2/c
31	80	50.3	4530	18	AA71253 Human ERBB2 gene.
32	80	50.3	4530	21	AA60815 Nucleotide sequenc
33	80	50.3	4530	22	AA19731 Human tyrosine kin
34	80	50.3	4530	24	AB235012 Human gene express
35	80	50.3	4530	24	ABV94128 Breast carcinoma r
36	80	50.3	4530	24	ABN85585 Human HER2-neu SEQ
37	80	50.3	4530	24	ABK83918 Human CDNA differe
38	80	50.3	4530	25	ACC50139 Breast cancer asso
39	80	50.3	4530	25	ABQ83856 Human Her2/Neu enc
40	80	50.3	9274	22	AAF24297 HER2 transgene pla
41	80	50.3	9274	24	AA43934 HER-2 transgene pl
42	80	50.3	9274	24	ABK14057 Human HER2 (ErBB2)
43	79	49.7	1115	21	AAF21778 Human breast and o
44	79	49.7	1755	24	AA322746 Human CDNA for the
45	79	49.7	1767	24	AA32744 Human CDNA for the

ALIGNMENTS

RESULT 1	
ABA92252	
ID	ABA92252 standard; CDNA; 2763 BP.
XX	
AC	ABA92252;
XX	
DT	17-JUN-2002 (first entry)
XX	
DE	Mouse Her-2/neu extracellular-phosphorylation domain fusion cDNA.
XX	
KW	Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
KW	receptor; mouse; gene therapy; gene; ss.
XX	
OS	Mus musculus.
XX	
FH	Key Location/Qualifiers
FT	CDS 1..2763

```

FT      /*tag= a
FT      /product= "Her-2/neu ECD-PD fusion"
PN      WO200212341-A2.
XX
PD      14-FEB-2002.
XX
PF      03-AUG-2001; 2001WO-US24283.
XX
PR      03-AUG-2000; 2000US-0632507.
XX
XX      (CORI-) CORIXA CORP.
PA      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX      Cheever MA, Gheysen D;
XX
DR      WPI; 2002-241743/29.
DR      P-PSDB; AAM51152.
XX
XX      Her-2/neu fusion protein for treating or preventing cancer by eliciting
PT      or enhancing an immune response to the protein, has Her-2/neu
PT      extracellular domain fused to Her-2/neu intracellular or
PT      phosphorylation domain -
XX
XX      Disclosure; Fig 23; 14lpp; English.
XX
XX      The present sequence is that of cDNA encoding a fusion between
CC      the extracellular domain (ECD) and the phosphorylation domain (PD)
CC      murine Her-2/neu oncoprotein (see AAM51151). The Her-2/neu gene is
CC      amplified and the oncoprotein is overexpressed in a variety of
CC      human cancers, including breast, ovarian, colon, lung and prostate
CC      cancer. Her-2/neu overexpression correlates with a poor prognosis
CC      in breast and ovarian cancers. The invention provides Her-2/neu
CC      fusion proteins, nucleic acids encoding them, viral vectors, and
CC      vaccines comprising the fusion proteins or nucleic acid molecules.
CC      In preferred fusion proteins, the ECD of Her-2/neu is fused to a
CC      Her-2/neu intracellular domain or PD (or its DeltapD fragment). An
CC      immune response to Her-2/neu protein is elicited or enhanced by
CC      administering the fusion protein in the form of a vaccine, or by
CC      transfecting cells of an animal ex vivo with a nucleic acid
CC      encoding the fusion protein, and delivering the transfectant cells
CC      to the animal. The fusion proteins, nucleic acids, and isolated
CC      cancer, especially breast, ovarian, colon, lung or prostate cancer
CC      in a patient. T cells that specifically react with a Her-2/neu
CC      fusion protein can be used to remove tumour cells from a sample in
CC      order to inhibit the development of cancer in a patient.
XX
XX      Sequence 2763 BP; 571 A; 855 C; 772 G; 565 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.:      1.16e-10      Length:      2763
XX      Score:          144.00      Matches:      25
XX      Percent Similarity: 100.00%      Conservative: 5
XX      Best Local Similarity: 83.33%      Mismatches: 0
XX      Query Match:      90.57%      Indels:      0
XX      DB:              24      Gaps:         0
XX
XX      JUNC_SEQ3_SEQ4 (1-30) x ABA92252 (1-2763)
XX
Oy      1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db      1918 GAACGAGGCTGCCAGCAGACGAGAGAGCCGCCAGTCTCTCAGAACGAGGACTTA 1977
Oy      21 GlyProAlaSerProLeuAspSerThrPhe 30
Db      1978 GGCCCTCCAGCCCTCATGGACAGCACCTTC 2007
XX
RESULT 2
ID      ABA92253
XX      ID ABA92253 standard; cDNA; 2781 BP.
XX
AC      ABA92253;

```

```

XX      17-JUN-2002 (first entry)
XX
DE      Her-2/neu extracellular-phosphorylation domain-Tcp0 fusion cDNA.
XX
KW      Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;
KW      receptor; Tcp0; mouse; gene therapy; gene; ss.
XX
OS      Chimeric - Mus musculus.
OS      Chimeric - Unidentified.
XX
FH      Key      Location/Qualifiers
FT      CDS      1..2781
FT      /*tag= a
FT      /product= "Her-2/neu ECD-PD-Tcp0 fusion"
XX
XX      WO200212341-A2.
XX
XX      14-FEB-2002.
XX
XX      03-AUG-2001; 2001WO-US24283.
XX
XX      03-AUG-2000; 2000US-0632507.
XX
XX      (CORI-) CORIXA CORP.
PA      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX      Cheever MA, Gheysen D;
XX
XX      WPI; 2002-241743/29.
XX      P-PSDB; AAM51153.
XX
XX      Her-2/neu fusion protein for treating or preventing cancer by eliciting
PT      or enhancing an immune response to the protein, has Her-2/neu
PT      extracellular domain fused to Her-2/neu intracellular or
PT      phosphorylation domain -
XX
XX      Disclosure; Fig 25; 14lpp; English.
XX
XX      The present sequence is that of cDNA encoding a fusion between
CC      the extracellular domain (ECD) and the phosphorylation domain (PD)
CC      murine Her-2/neu oncoprotein (see AAM51151) plus a C-terminal
CC      Tcp0 motif that improves immunogenicity. The Her-2/neu gene is
CC      amplified and the oncoprotein is overexpressed in a variety of
CC      human cancers, including breast, ovarian, colon, lung and prostate
CC      cancer. Her-2/neu overexpression correlates with a poor prognosis
CC      in breast and ovarian cancers. The invention provides Her-2/neu
CC      fusion proteins, nucleic acids encoding them, viral vectors, and
CC      vaccines comprising the fusion proteins or nucleic acid molecules.
CC      In preferred fusion proteins, the ECD of Her-2/neu is fused to a
CC      Her-2/neu intracellular domain or PD (or its DeltapD fragment). An
CC      immune response to Her-2/neu protein is elicited or enhanced by
CC      administering the fusion protein in the form of a vaccine, or by
CC      transfecting cells of an animal ex vivo with a nucleic acid
CC      encoding the fusion protein, and delivering the transfectant cells
CC      to the animal. The fusion proteins, nucleic acids, and isolated
CC      specific T-cells are useful for inhibiting the development of a
CC      cancer, especially breast, ovarian, colon, lung or prostate cancer
CC      in a patient. T cells that specifically react with a Her-2/neu
CC      fusion protein can be used to remove tumour cells from a sample in
CC      order to inhibit the development of cancer in a patient.
XX
XX      Sequence 2781 BP; 574 A; 859 C; 779 G; 569 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.:      1.17e-10      Length:      2781
XX      Score:          144.00      Matches:      25
XX      Percent Similarity: 100.00%      Conservative: 5
XX      Best Local Similarity: 83.33%      Mismatches: 0
XX      Query Match:      90.57%      Indels:      0
XX      DB:              24      Gaps:         0
XX
XX      JUNC_SEQ3_SEQ4 (1-30) x ABA92253 (1-2781)

```

QY 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20  
Db 1918 GAACGAGGCTGCCAGCAGACAGAGAGCCAGCCAGTCTCAGAACGAGACTTA 1977  
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30  
Db 1978 GGCCCTCCAGCCCATGGAGACAGCACCCTC 2007

RESULT 3  
AAT72725  
ID AAT72725 standard; cDNA; 2385 BP.  
AC AAT72725;  
DT 17-SEP-1997 (first entry)  
DE Her2-GM-CSF immunostimulant fusion protein DNA.  
KW Her2-GM-CSF; granulocyte macrophage colony stimulating factor;  
KW growth factor receptor; oncogene; immunostimulant; cancer;  
KW therapy; ss.  
OS Homo sapiens.

XX Key Location/Qualifiers  
FH 11..2359  
FT CDS /\*tag= a  
FT /product= GM-CSF-Her2 fusion protein  
FT mRNA 11..1969  
FT /\*tag= b  
FT /product= Her2  
FT mRNA 1970..1975  
FT /\*tag= c  
FT /product= Leu-Glu linker  
FT mRNA 1976..2359  
FT /\*tag= d  
FT /product= GM-CSF

XX WO9724438-A1.  
XX 10-JUL-1997.  
XX 23-DEC-1996; 96WO-US20241.  
XX 28-DEC-1995; 95US-0579823.  
XX (ACTI-) ACTIVATED CELL THERAPY INC.  
XX Laus R, Ruegg CL, Wu H;  
XX WPI; 1997-363674/33.  
XX P-PSDB; AAW19764.

XX Potent APC that activates T-cells to give multivalent cellular  
XX immune response - can also induce a cytotoxic T-cell response in a  
XX vertebrate subject  
XX Disclosure; Fig 8; 45pp; English.

XX A nucleic acid molecule (AAW72725) codes for a fusion protein  
XX (AAW19764) comprising granulocyte-macrophage colony stimulating  
XX factor (GM-CSF) and Her2, a growth factor receptor that is  
XX over-expressed in breast and ovarian cancer cells. It was  
XX prepd. by PCR amplification of Her2 cDNA from a breast cancer  
XX cell line and fusion to GM-CSF cDNA. Fusion expression vectors can  
XX be used to transfect mammalian and insect cells. The Her2-GM-CSF  
XX fusion protein is used to generate anti-Her2 immunity. Tumour  
XX cells are eliminated by cytotoxic T lymphocytes activated in vivo  
XX or in vitro by exposure to antigen-presenting cells exposed to the  
XX fusion protein.

XX Sequence 2385 BP; 488 A; 780 C; 677 G; 440 T; 0 other;

Alignment Scores:  
Pred. No.: 0.0382 Length: 2385  
Score: 84.00 Matches: 21  
Percent Similarity: 70.00% Conservative: 0  
Best Local Similarity: 70.00% Mismatches: 5  
Query Match: 52.83% Indels: 4  
DB: 18 Gaps: 2  
JUNC\_SEQ3\_SEQ4 (1-30) x AAT72725 (1-2385)  
QY 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20  
Db 1925 GACAAGGCTGCCCGCCGAGAGAGCCAGCCCTCTGACGTCCTCGAG----- 1975  
QY 21 GlyProAla---SerProLeuAspSerThr 29  
Db 1976 GCACCGCGCGCTCGCCCGAGCCGAGCACA 2005  
RESULT 4  
AA14815  
ID AA14815 standard; cDNA; 69 BP.  
AC AA14815;  
XX 08-AUG-2000 (first entry)  
DE DNA encoding a wild type erbB-2 receptor protein fragment.  
XX SPLICE erbB-2 receptor protein; cell transformation disorder; cancer;  
KW tumor cell proliferation; tissue degeneration; arthropathy;  
KW bone resorption; inflammatory disease; degenerative disorder;  
KW wound healing; ss.  
XX Homo sapiens.  
XX WO200020579-A1.  
XX 13-APR-2000.  
XX 01-OCT-1999; 99WO-CA00912.  
XX 02-OCT-1998; 98US-0165192.  
XX (UYMC-) UNIV MCMASTER.  
XX Muller WJ, Siegel PM;  
XX WPI; 2000-303768/26.  
XX P-PSDB; AAY84681.  
XX Nucleic acid encoding an erbB 2 receptor protein designated SPLICE  
XX erbB-2, inhibitors of the protein are useful for treatment of cancer -  
XX Example 2; Fig 3B; 60pp; English.  
XX The present sequence encodes a peptide of wild type erbB-2 receptor  
XX protein, where the splicing event occurs. An alternatively spliced  
XX variant of erbB-2 receptor protein, designated SPLICE erbB-2 receptor  
XX protein, has an in-frame deletion of 16 amino acids, 2 of which are  
XX conserved cysteine residues, compared to the unspliced protein. The  
XX erbB-2 polynucleotide is used to construct probes for detecting  
XX disorders of cell transformation such as cancer. Antibodies to the  
XX protein may be used to detect SPLICE erbB-2 in a sample. Agents  
XX (e.g. antisense oligonucleotides) which inhibit the expression of  
XX SPLICE erbB-2 are useful for reducing tumor cell proliferation and  
XX treating cancer. Substances which stimulate SPLICE erbB-2 are useful  
XX for treating conditions involving damaged cells including conditions in  
XX which degeneration of tissue occurs, such as arthropathy, bone  
XX resorption, inflammatory diseases, degenerative disorders of the  
XX central nervous system and wound healing.  
XX Sequence 69 BP; 13 A; 26 C; 20 G; 10 T; 0 other;

```

Alignment Scores:
Pred. No.:      0.00195      Length:      69
Score:          80.00      Matches:      15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      50.31%      Indels:      0
DB:              21      Gaps:      0

JUNC_SEQ3_SEQ4 (1-30) x AAA14815 (1-69)

QY      1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db      25 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCTCTGACGTC 69

RESULT 5
AAAG62125
ID      AAA62125 standard; RNA; 201 BP.
XX
AC      AAA62125;
XX
DT      20-JUN-2001 (first entry)
XX
DE      Her2 normal form transcript.
XX
KW      Her2; disease prediction; breast cancer; human; tyrosine kinase; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      stem_loop      Location/Qualifiers
FT      8..18      /*tag= a
FT      stem_loop      37..56      /*tag= b
FT      misc_binding      58..59      /*tag= c
FT      /*bound_moiety= "Her2 normal form transcript"
FT      /*note= "Forms double-stranded region with bases 74-75
FT      of Her2 normal form transcript"
FT      stem_loop      62..71      /*tag= d
FT      misc_binding      74..75      /*tag= e
FT      /*bound_moiety= "Her2 normal form transcript"
FT      /*note= "Forms double-stranded region with bases 58-59
FT      of Her2 normal form transcript"
FT      misc_binding      78..81      /*tag= f
FT      /*bound_moiety= "Her2 normal form transcript"
FT      /*note= "Forms double-stranded region with bases 121-124
FT      of Her2 normal form transcript"
FT      misc_binding      84..90      /*tag= g
FT      /*bound_moiety= "Her2 normal form transcript bases"
FT      /*note= "Forms double-stranded region with bases 111-117
FT      of Her2 normal form transcript"
FT      stem_loop      94..107      /*tag= h
FT      misc_binding      111..117      /*tag= i
FT      /*bound_moiety= "her2 normal form transcript"
FT      /*note= "Forms double-stranded region with bases 84-90
FT      of Her2 normal form transcript"
FT      misc_binding      121..124      /*tag= j
FT      /*bound_moiety= "Her2 normal form transcript"
FT      /*note= "Forms double-stranded region with bases 78-81
FT      of Her2 normal form transcript"
FT      stem_loop      126..132      /*tag= k
FT      misc_binding      137..152      /*tag= l
FT      /*bound_moiety= "Her2 normal form transcript bases"

/note= "Forms double-stranded region with bases 172-187
of Her2 normal form transcript"
/*tag= m
/*tag= n
/*bound_moiety= "her2 normal form transcript bases"
/note= "Forms double-stranded region with bases 137-152
of Her2 normal form transcript"

stem_loop
misc_binding
/*tag= m
/*tag= n
/*bound_moiety= "her2 normal form transcript bases"
/note= "Forms double-stranded region with bases 137-152
of Her2 normal form transcript"

WO2000311110-A1.
02-JUN-2000.
22-NOV-1999; 99WO-US27710.
25-NOV-1998; 98US-0110024.
25-NOV-1998; 98US-0200355.
(ISIS-) ISIS PHARM INC.
Ecker DJ;
WPI; 2000-400027/34.

Identifying a target nucleic acid sequence predictive of preselected
disease states such as a cancerous state, by comparing members of a set
of mRNA molecules, from a common gene, containing different sequences
and structures -
Example 2; Fig 3; 38pp; English.

The HER2 proto-oncogene encodes a tyrosine kinase receptor, which
functions in cellular signal transduction. The HER2 protein is
implicated in breast cancer. The HER2 receptor mRNA exists in at least
two forms: a normal form and a truncated form. The present sequence is
the normal form RNA of HER2. The truncated form RNA sequence of HER2
(AAA62126) encodes a protein associated with increased resistance to the
growth inhibiting effects of a monoclonal antibody, Herceptin, used in
cancer treatment. The truncated form of the transcript contains unique
structures not found in the normal form. The present sequence is
predictive of Herceptin-resistant cancer and detection of the present
sequence may therefore be used as a method of diagnosing breast cancer.
Other diseases which may be identified by using a similar method to
detect other RNA molecules are hyperproliferative conditions, Lupus
erythematosus, psoriasis, inflammation, cardiovascular disease, pain,
arthritis, obesity, trauma, Huntington's disease or neurological
disorders.

Sequence 201 BP; 34 A; 67 C; 56 G; 44 U; 0 other;

Alignment Scores:
Pred. No.:      0.00714      Length:      201
Score:          80.00      Matches:      15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      50.31%      Indels:      0
DB:              21      Gaps:      0

JUNC_SEQ3_SEQ4 (1-30) x AAA62125 (1-201)

QY      1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db      116 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCTCTGACGTC 160

RESULT 6
AAV21727
ID      AAV21727 standard; cDNA; 2125 BP.
XX
AC      AAV21727;
XX
DT      17-AUG-1998 (first entry)
XX

```

```

DE Humanised vector pITL-hHER/neu.
KW Vector; vaccine; tumour; antigen; plasmid pITL-hHER/neu;
KW human; HER-2/neu; C-erbB-2; breast cancer; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Escherichia coli.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 13..921
FT /tag= a
FT /product= human HER-2/neu
FT polyA_site 922..1181
FT /tag= b
FT /note= "Combined splice and polyA sequences"
FT CDS 1195..1401
FT /tag= c
FT /note= "SupF gene"
FT misc_feature 1412..1864
FT /tag= d
FT /note= "ColE1 origin of replication"
FT promoter 1877..2125
FT /tag= e
FT /note= "RANTES promoter"
FT
FT
XX WO9806863-A1.
XX
XX 19-FEB-1998.
XX
XX 14-AUG-1997; 97WO-US14306.
XX
XX 14-AUG-1996; 96US-0023931.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson EL, Nelson PJ;
XX WPI; 1998-159552/14.
XX
XX Humanised polynucleotide vectors - comprising human derived promoter
XX and sequence acceptance site, used for the production of vaccines
XX
XX Example 7; Page 41-42; 125pp; English.
XX
XX Plasmid pITL-hHER2/neu comprises base vector pITL (see AAV211724) and
XX a human HER-2/neu nucleic sequence. pITL-hHER2/neu was used to
XX evaluate the toxicity of anti-tumour vaccination in rats, and in
XX phase I and phase II trials to evaluate polynucleotide vaccination
XX in advanced breast cancer. Novel humanised vectors, which can be
XX based on pITL, comprise a human-derived promoter or mammalian
XX homologue which is functional in mammalian target tissue and cells
XX and a sequence acceptance site which accepts cDNA products from
XX RT-PCR cloning. The vectors are non-replicating in mammalian cells
XX but are capable of extended stable expression in the target
XX sequence, generating an immune response in immunised individuals.
XX The vectors selectively elicit immune responses to the target
XX sequences with little or no immune response to the other components
XX of the vectors.
XX
XX Sequence 2125 BP; 449 A; 650 C; 598 G; 428 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 0.124 Length: 2125
XX Score: 80.00 Matches: 15
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 50.31% Indels: 0
XX DB: 19 Gaps: 0
XX
XX JUNC_SEQ3_SEQ4 (1-30) x AAV21727 (1-2125)
XX
XX 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

```

---

```

Db 112 GACAAGGGCTGCCCGCCGAGAGAGAGCCGCTGTGACGTCC 156
RESULT 7
AAZ50586
ID AAZ50586 standard; DNA; 2871 BP.
XX
AC AAZ50586;
XX
DT 23-MAY-2000 (first entry)
XX
DE DC8scFv-erbB2EC fusion construct containing tetramerisation domain.
DE
DE DC8 scFv; single-chain variable fragment; erbB2EC; extracellular domain;
DE human; fusion construct; tetramerisation domain; constant domain;
DE heteroinhibitor; multifunctional compound; melanoma; sarcoma;
DE immunoglobulin; cytostatic; immunostimulatory; antileukaemia; diagnosis;
DE antiproliferative; prevention; treatment; malignant; haematopoietic cell;
DE lymphoma; leukaemia; solid tumour; carcinoma; ds.
XX
OS Chimeric - Unidentified.
OS Chimeric - Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 10..2865
XX /tag= a
XX /product= "DC8scFv-erbB2EC fusion protein"
XX sig_peptide 10..66
XX /tag= b
XX /label= Leader_sequence
XX misc_feature 67..390
XX /tag= c
XX /label= DC8scFv_light_chain_variable_region
XX misc_feature 391..435
XX /tag= d
XX /label= Glycine-Serine-linker_DNA
XX misc_feature 436..771
XX /tag= e
XX /label= DC8scFv_heavy_chain_variable_region
XX misc_feature 775..807
XX /tag= f
XX /note= "5' end of human IgG3 upper hinge region
XX with 3 additional nucleotides"
XX misc_feature 808..924
XX /tag= g
XX /label= Human_p53_tetramerisation_domain
XX misc_feature 925..945
XX /tag= h
XX /label= Short_peptide_linker
XX misc_feature 946..2844
XX /tag= i
XX /label= erbB2EC_domain
XX misc_feature 2845..2862
XX /tag= j
XX /label= His_tag
XX
XX WO200006605-A2.
XX
XX 10-FEB-2000.
XX
XX 28-JUL-1999; 99WO-EP05416.
XX
XX 28-JUL-1998; 98EP-0114082.
XX (MICR-) MICROMET GES BIOMEDIZINISCHE FORSCHUNG.
XX Kufer P, Dreier T, Baeuerle PA, Borschert K, Zettl F;
XX WPI; 2000-195265/17.
XX P-PSDB; AAY44993.
XX
XX New multifunctional compounds useful for preventing and/or treating
XX malignant cell growth and for detection and diagnosis.

```

XX Example 9; Fig 49; 166pp; English.

PS The patent discloses heteroinibodies which are multifunctional compounds

XX producible in a mammalian host cell as a secretable and fully functional

CC heterodimer of two polypeptide chains, where one of the polypeptide

CC chains comprises, a CHI-domain (constant domain of an immunoglobulin

CC heavy chain) and the other chain comprises CL-domain (constant domain of

CC an immunoglobulin light chain). The polypeptide chains further comprise,

CC fused to the constant domains at least two (poly)peptides having

CC different receptor or ligand functions, where further at least two of the

CC different (poly)peptides lack an intrinsic affinity for one another and

CC are linked via the constant domains. The heteroinibodies have

CC cytostatic, immunostimulatory, antileukaemia and antiproliferative

CC activities. These compounds can be used for diagnosing, preventing and

CC treating malignant cell growth related to malignancies of haematopoietic

CC cells e.g. lymphomas and leukemias, or to solid tumours e.g. carcinomas,

CC melanomas and sarcomas.

CC The present sequence is a fusion construct comprising DC8

CC single-chain Fv (scFv) fragment at the N-terminus, extracellular

CC domain of human erbB2 at the C-terminus and a tetramerisation

CC domain between them. This construct was prepared to find out whether

CC an oligomerisation domain characterised in bacterial expression system

CC is applicable for expression of fully functional and secretable

CC recombinant protein in mammalian host cells. This tetrameric construct

CC was not expressed as secretable and fully functional protein

CC in mammalian cells. Hence general applicability of the tetramerisation

CC domain for oligomerisation strategies in mammalian cells was ruled out.

XX

SQ Sequence 2871 BP; 598 A; 868 C; 834 G; 571 T; 0 other;

Alignment Scores:

Pred. No.:	0.179	Length:	2871
Score:	80.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	50.31%	Indels:	0
DB:	21	Gaps:	0

JUNC\_SEQ3\_SEQ4 (1-30) x AAZ50586 (1-2871)

OY 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

DB 2797 GACAAGGGCTGCCCGCGCAGAGAGAGCCAGCCCTCTGACGTCC 2841

RESULT 8

AAA89736

ID AAA89736 standard; DNA; 3600 BP.

XX

AC AAA89736;

XX

DT 12-JAN-2001 (first entry)

XX

DE Human HER-2/neu coding sequence.

XX

KW Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;

KW breast cancer; prostate cancer; ovarian cancer; lung cancer;

KW colon cancer; ds.

XX

OS Homo sapiens.

XX

Key Location/Qualifiers

FT CDS 1..3600

FT /tag= a

FT /product= "HER-2/neu protein"

XX

PN WO200044899-A1.

XX

PD 03-AUG-2000.

XX

PF 28-JAN-2000; 2000WO-US02164.

XX

PR 29-JAN-1999; 99US-0117976.

XX (CORI-) CORIXA CORP.

PA (SMIK ) SMITHLINE BEECHAM.

XX

PI Cheever MA, Gheysen D;

XX

DR WPI; 2000-505976/45.

XX

DR P-PSDB; AAB21198, AAB21208.

XX

PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins

PT useful for vaccinating against breast, ovarian, colon, lung and

PT prostate cancers -

XX

PS Disclosure; Fig 15; 128pp; English.

XX

CC The present sequence encodes the human HER-2/neu protein. HER-2/neu is

CC a member of the tyrosine kinase family of receptor-like glycoproteins

CC and shows homology to the epidermal growth factor receptor (EGFR). It

CC probably plays a part in cell growth and/or differentiation. The

CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising

CC a HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation

CC domain may be used to treat or prevent cancer by eliciting or

CC enhancing an immune response to the HER-2/neu protein. It may be used

CC to treat malignancies such as breast, ovarian, colon, lung and

CC prostate cancers, and may be used as an antigen to vaccinate against

CC these neoplasias.

XX

SQ Sequence 3600 BP; 723 A; 1108 C; 1075 G; 694 T; 0 other;

Alignment Scores:

Pred. No.:	0.235	Length:	3600
Score:	80.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	50.31%	Indels:	0
DB:	21	Gaps:	0

JUNC\_SEQ3\_SEQ4 (1-30) x AAA89736 (1-3600)

OY 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

DB 1915 GACAAGGGCTGCCCGCGCAGAGAGCCAGCCCTCTGACGTCC 1959

RESULT 9

ABK86207

ID ABK86207 standard; cDNA; 3678 BP.

XX

AC ABK86207;

XX

DT 24-SEP-2002 (first entry)

XX

DE cDNA encoding human breast cancer antigen, Her2 variant.

XX

KW Human; Her2; cytostatic; antiviral; immunostimulant;

KW cell-mediated immune response; tumour; breast cancer;

KW virus infection; prostate cancer; colorectal cancer; pancreatic cancer;

KW lymphoma; leukaemia; hepatitisvirus; lentivirus; herpesvirus;

KW human immunodeficiency virus; HIV; flavivirus; pestivirus; gene; ss.

XX

OS Homo sapiens.

XX

Key Location/Qualifiers

FT CDS 7..3678

FT /tag= a

FT /product= "Breast cancer antigen Her2 variant"

XX

PN WO200240059-A2.

XX

PD 23-MAY-2002.

XX

PF 01-NOV-2001; 2001WO-US45626.

XX

PR 01-NOV-2000; 2000US-0704232.



XX (AMBI-) AMERICAN FOUND BIOLOGICAL RES INC.  
 PA (MINC/) MINCHEFF M S.  
 PA (LOUK/) LOUKINOV D I.  
 PA (ZOUB/) ZOUBAK S.  
 XX Mincheff MS, Loukinov DI, Zoubak S;  
 XX WPI: 2002-527524/56.  
 DR P-PSDB; AAU98923.  
 XX  
 PT Inducing a cell-mediated immune response against a target antigen,  
 PT reducing undesired cells and stimulating presentation of an antigen by  
 PT a cell, comprises administering a polynucleotide encoding a variant of  
 PT an antigen -  
 XX  
 XX Disclosure; Page 128-134; 146pp; English.  
 PS  
 XX The invention relates to a method of inducing a cell-mediated immune  
 CC response against a cell comprising a target antigen (I) in a subject,  
 CC treating a subject having undesired cells for example tumour cells  
 CC or virally infected cells (C), reducing the number of (C) in a subject,  
 CC and stimulating presentation of (I) by a cell. This is done by  
 CC administering a polynucleotide (II) encoding a variant of (I), so that  
 CC (II) expressed in a cell and cell-mediated immune response is induced.  
 CC The method can be used to treat prostate cancer, breast cancer,  
 CC colorectal cancer and pancreatic cancer, as well as lymphomas and  
 CC leukemias. The method is also useful in treating chronic viral  
 CC infections such as those caused by hepatitis viruses, lentiviruses  
 CC (including human immunodeficiency virus (HIV)), herpesviruses and the  
 CC flaviviruses and pestiviruses. The present sequence represents the coding  
 CC sequence of human breast cancer antigen, Her2 variant, used as a target  
 CC antigen in the method of the invention.  
 XX  
 XX SQ Sequence 3678 BP; 746 A; 1137 C; 1091 G; 704 T; 0 other;

Alignment Scores:  
 Pred. No.: 0.241 Length: 3678  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.31% Indels: 0  
 DB: 24 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x ABR86207 (1-3678)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
 DB 1825 GACAGGGGTGCCCGCCGAGCAGAGAGCCCTCTGACGTC 1869

RESULT 10

AAAT40739  
 ID AAAT40739 standard; cDNA; 3768 BP.

XX AAAT40739;

XX 01-JAN-1997 (first entry)  
 XX  
 XX DE HER-2/neu oncogene.

XX HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;  
 KW breast cancer; ovary cancer; colon cancer; lung cancer;  
 KW prostate cancer; genetic immunisation; tumour; vaccine; vector;  
 KW ss.  
 XX Homo sapiens.

XX Key Location/Qualifiers  
 FH CDS 1..3765  
 FT /\*tag= b  
 FT /note= "nucleotides 2026-3765 (claim 1) code for  
 FT HER-2/neu intracellular domain"  
 XX

PN WO9630514-A1.  
 XX  
 PD 03-OCT-1996.  
 XX  
 XX PF 28-MAR-1996; 96WO-US01689.  
 XX  
 XX PR 31-MAR-1995; 95US-0414417.  
 XX  
 XX PA (UNIW ) UNIV WASHINGTON.  
 XX  
 XX PI Cheever MA, Disis ML;  
 XX WPI: 1996-455361/45.  
 DR P-PSDB; AAU01111.  
 XX  
 PT DNA encoding HER-2-neu poly:peptide(s) - used for prevention or  
 PT treatment of malignancies with which the HER-2/neu oncogene is  
 PT associated  
 PT  
 XX Claim 1; Page 49-56; 71pp; English.  
 XX  
 CC Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or  
 CC c-erbB2) protein (AAU01111). The oncogene is overexpressed in various  
 CC cancers, including breast, ovarian, colon, lung and prostate, and  
 CC appears to induce malignancies through quantitative mechanisms that  
 CC result from increased or deregulated expression of an essentially  
 CC normal gene product. Nucleotides 2026-3765 of the cDNA sequence  
 CC code for the intracellular domain (lys676-val1255) of the HER-2/neu  
 CC protein, which is useful for immunisation against malignancy.  
 CC Nucleic acids can be used to direct expression of the intracellular  
 CC domain in transformed host cells, or are used, alone or in a viral  
 CC vector, for genetic immunisation of an animal.  
 XX  
 XX SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:  
 Pred. No.: 0.248 Length: 3768  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.31% Indels: 0  
 DB: 17 Gaps: 0

JUNC\_SEQ3\_SEQ4 (1-30) x AAT40739 (1-3768)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
 DB 1915 GACAGGGGTGCCCGCCGAGCAGAGAGCCCTCTGACGTC 1959

RESULT 11

AAAX01912  
 ID AAAX01912 standard; DNA; 3768 BP.

XX AAAX01912;

XX 21-APR-1999 (first entry)  
 XX  
 XX DE Human HER-2/neu oncogene DNA.

XX HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;  
 KW malignancy; treatment; tumour; ss.  
 KW Homo sapiens.

XX Key Location/Qualifiers  
 FH CDS 1..3768  
 FT /\*tag= a  
 FT /product= "HER-2/neu"  
 FT /note= "oncogene"  
 FT misc\_feature 2026..3765  
 FT /\*tag= b  
 FT /note= "region which elicits immune response"  
 XX

```

PN US5869445-A.
XX
PD 09-FEB-1999.
XX
PF 01-APR-1996; 96US-0625101.
XX
PR 01-APR-1996; 96US-0625101.
PR 17-MAR-1993; 93US-0033644.
PR 12-AUG-1993; 93US-0106112.
PR 31-MAR-1995; 95US-0414417.
XX
PA (UNIW ) UNIV WASHINGTON.
XX
PI Cheever MA, Disis ML;
XX
DR WPI; 1999-152835/13.
DR P-PSDB; AA92406.
XX
PT Use of HER-2/neu polypeptides - for eliciting an immune response to
PT an HER-2/neu associated malignancy, particularly for treating or
PT preventing tumours
XX
PS Claim 1a; Column 23-32; 26pp; English.
XX
CC This sequence encodes the human HER-2/neu oncogene protein. A fragment
CC of this protein is used in a method for eliciting or enhancing an immune
CC response to HER-2/neu protein. The polypeptide can stimulate T cells and
CC B cells to produce an immune response to the HER-2/neu protein. The
CC method can be used for immunisation against a malignancy in which the
CC HER-2/neu oncogene is associated and in the treatment of an existing
CC tumour, or to prevent tumour occurrence or reoccurrence.
XX
SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 0.248 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 20 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x AA01912 (1-3768)
QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
DB 1915 GACRAGGGCTGCCCGCGAGCAGAGACGACGACCCCTCTGACGTCC 1959

RESULT 12
AAA09455
ID AAA09455 standard; DNA; 3768 BP.
XX
AC AAA09455;
XX
DT 10-AUG-2000 (first entry)
XX
DE Human heregulin 2 (Her2) coding sequence.
XX
KW Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
KW self-protein; cell-associated peptide antigen; foreign epitope;
KW cancer; breast cancer; prostate cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200020027-A2.
XX
PD 13-APR-2000.
XX
PF 05-OCT-1999; 99WO-DK00525.
XX
PR 05-OCT-1998; 98DK-0001261.
PR 20-OCT-1998; 98US-0105011.
XX

```

```

PA (MEBI-) M & E BIOTECH AS.
XX
PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
PI Gautam A, Birk P, Karlsson G;
XX
DR WPI; 2000-349917/30.
DR P-PSDB; AAY92620.
XX
PT Inducing immune responses to weakly immunogenic, tumor associated
PT peptide antigens for the treatment of breast and prostate cancer
XX
PS Claim 62; Page 187-193; 220pp; English.
XX
CC The claims detail a method for inducing immune responses against weakly
CC immunogenic cell-associated peptide antigens (PA) such as those
CC associated with cancers (i.e. self-proteins), for example, human
CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
CC simultaneous presentation by antigen producing cells (APCs) of the
CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
CC group derived from the PA and/or at least 1 B-cell group derived from the
CC cell-associated PA; and (2) at least 1 first T helper cell group which is
CC foreign to the animal. Analogues of human PSM, human Her2 and
CC human/murine FGF8b comprising a substantial part of all known and
CC predicted CTL and B-cell epitopes of the respective PA and including at
CC least one foreign T helper epitope are also claimed. The method is used
CC to treat prostate, prostate/breast or breast cancer when the PA is human
CC PSM, FGF8b and Her2, respectively.
XX
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 0.248 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.31% Indels: 0
DB: 21 Gaps: 0

JUNC_SEQ3_SEQ4 (1-30) x AAA09455 (1-3768)
QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
DB 1915 GACRAGGGCTGCCCGCGAGCAGAGACGACGACCCCTCTGACGTCC 1959

RESULT 13
AAH23392
ID AAH23392 standard; DNA; 3768 BP.
XX
AC AAH23392;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human HER-2/neu protein encoding DNA.
XX
KW Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
KW oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3768
FT /tag= a
FT /product= "HER-2/neu protein"
XX
PN WO200153463-A2.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US01850.
PR 21-JAN-2000; 2000US-0177545.
XX

```



CC (ds) structure of not more than 49 consecutive nucleotides (nt), where  
CC at least a segment of one strand of the ds structure is complementary  
CC with the target gene and the cells are treated with interferon before  
CC introduction of dsRNA1. The method is used to inhibit expression of  
CC target genes, particularly oncogenes, cytokine genes, id (not defined)  
CC protein genes; developmental or prion genes, or genes expressed in  
CC pathogenic organisms (particularly plasmidia) or in viruses or viroids  
CC (pathogenic in humans, animals or plants). Treating the cells with  
CC interferon greatly increases the extent to which dsRNA can inhibit  
CC expression of the target genes, and the effect is even greater when dsRNA  
CC are modified to increase their stability. ABX09936-ABX10075 represent  
CC gene fragments used to illustrate the method of the invention.  
XX

SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:			
Pred. No.:	0.248	Length:	3768
Score:	80.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	50.31%	Indels:	0
DB:	24	Gaps:	0
JUNC_SEQ3_SEQ4 (1-30) x ABX09987 (1-3768)			
QY	1 AsPLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15		
Db	 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTCC 1959		

Search completed: October 15, 2003, 23:48:38  
Job time : 31.4941 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:09:13 ; Search time 432.026 Seconds  
(without alignments)  
2840.777 Million cell updates/sec

Title: JUNC\_SEQ3\_SEQ4

Perfect score: 159

Sequence: 1 DKGCPAEQRASPLTSQNDLGPA SPLDSTF 30

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlip  
-Q=/cgn2\_1/USPTO.spool\_p/HOLLERAN480/runat\_15102003\_131912\_20536/app\_query.fasta\_1.4685  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 @CGN1\_1\_22724\_runat\_15102003\_131912\_20536 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_ov.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_in.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_mu.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	144	90.6	2763	6	AX380942	AX380942 Sequence
2	144	90.6	2781	6	AX380944	AX380944 Sequence
C 3	85.5	53.8	186134	10	AL591390	AL591390 Mouse DNA
C 4	85.5	53.8	214019	2	AC064803	AC064803 Mus muscu
5	85	53.5	13450	9	AB096612	AB096612 Homo sapi
6	85	53.5	16572	9	AB096614	AB096614 Homo sapi
7	85	53.5	20271	9	AB096613	AB096613 Homo sapi
8	85	53.5	30837	9	AY208911	AY208911 Homo sapi
9	85	53.5	161815	9	AC079199	AC079199 Homo sapi
10	85	53.5	168585	9	AC040933	AC040933 Homo sapi
11	85	53.5	198008	2	AC142197	AC142197 Homo sapi
12	84	52.8	2385	6	AR082744	AR082744 Sequence
13	84	52.8	2385	6	AR099963	AR099963 Sequence
14	84	52.8	2385	6	AR143949	AR143949 Sequence
15	80	50.3	153	6	AX022984	AX022984 Sequence
16	80	50.3	201	6	AR229723	AR229723 Sequence
17	80	50.3	281	6	AX023363	AX023363 Sequence
18	80	50.3	3678	6	AX051114	AX051114 Sequence
19	80	50.3	3768	6	AR034479	AR034479 Sequence
20	80	50.3	3768	6	AX060704	AX060704 Sequence
21	80	50.3	3768	6	AX201817	AX201817 Sequence
22	80	50.3	3768	6	AX380923	AX380923 Sequence
23	80	50.3	3768	6	AX384604	AX384604 Sequence
24	80	50.3	3768	6	AX465456	AX465456 Sequence
25	80	50.3	3768	6	AX467229	AX467229 Sequence
26	80	50.3	3768	6	AX481438	AX481438 Sequence
27	80	50.3	4473	6	AR080259	AR080259 Sequence
28	80	50.3	4473	6	AR167390	AR167390 Sequence
29	80	50.3	4473	9	HSEB2R	X03363 Human c-erb
30	80	50.3	4530	6	AR202597	AR202597 Sequence
31	80	50.3	4530	6	AR283481	AR283481 Sequence
32	80	50.3	4530	6	AX282577	AX282577 Sequence
33	80	50.3	4530	6	AX587649	AX587649 Sequence
34	80	50.3	4530	6	AX644071	AX644071 Sequence
35	80	50.3	4530	6	BD005474	BD005474 Cellular
36	80	50.3	4530	6	I21124	I21124 Sequence 9
37	80	50.3	4530	6	I59745	I59745 Sequence 9
38	80	50.3	4530	9	HUMHER2A	M11730 Human tyros
39	80	50.3	9274	6	AX060703	AX060703 Sequence
40	79	49.7	1755	6	AX384609	AX384609 Sequence
41	79	49.7	1757	6	AX384607	AX384607 Sequence
42	79	49.7	1773	6	AX384610	AX384610 Sequence
43	79	49.7	1806	6	AX384608	AX384608 Sequence
44	79	49.7	3780	4	AB008451	AB008451 Canis fam
45	76	47.8	4062	10	HANNEU	D16295 Mesocricetu

#### ALIGNMENTS

RESULT 1

```

AX380942
LOCUS AX380942 2763 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 28 from Patent WO0212341.
ACCESSION AX380942
VERSION AX380942.1 GI:19575786
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
          artificial sequences.
REFERENCE
  1 Cheever,M.A. and Gheysen,D.
  TITLE Her-2/neu fusion proteins
  JOURNAL Patent: WO 0212341-A 28 14-FEB-2002;
  FEATURES
    Location/Qualifiers
      source
        1..2763
        /organism="synthetic construct"
        /mol_type="genomic DNA"
        /db_xref="taxon:32630"
        /note="mouse ECD-PD fusion protein cDNA"
BASE COUNT 571 a 855 c 772 g 565 t
ORIGIN
Alignment Scores:
Pred. No.:      8.97e-10      Length:      2763
Score:          144.00      Matches:      25
Percent Similarity: 100.00%      Conservative: 5
Best Local Similarity: 83.33%      Mismatches: 0
Query Match:      90.57%      Indels:      0
DB:              6          Gaps:      0

JUNC_SEQ3_SEQ4 (1-30) x AX380942 (1-2763)
QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCAGCAGACAGACGACCCAGTCAGCTCTCAGACGAGACTTA 1977
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 1978 GGCCTCCAGCCCATGGACGACACCTTC 2007
Db 1978 GGCCTCCAGCCCATGGACGACACCTTC 2007

RESULT 2
AX380944
LOCUS AX380944 2781 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 30 from Patent WO0212341.
ACCESSION AX380944
VERSION AX380944.1 GI:19575787
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
          artificial sequences.
REFERENCE
  1 Cheever,M.A. and Gheysen,D.
  TITLE Her-2/neu fusion proteins
  JOURNAL Patent: WO 0212341-A 30 14-FEB-2002;
  FEATURES
    Location/Qualifiers
      source
        1..2781
        /organism="synthetic construct"
        /mol_type="genomic DNA"
        /db_xref="taxon:32630"
        /note="mouse ECD-PD-Tcp0 fusion protein cDNA"
BASE COUNT 574 a 859 c 779 g 569 t
ORIGIN
Alignment Scores:
Pred. No.:      9.03e-10      Length:      2781
Score:          144.00      Matches:      25
Percent Similarity: 100.00%      Conservative: 5
Best Local Similarity: 83.33%      Mismatches: 0
Query Match:      90.57%      Indels:      0
DB:              6          Gaps:      0

```

```

JUNC_SEQ3_SEQ4 (1-30) x AX380944 (1-2781)
QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1918 GAACGAGGCTGCCAGCAGACAGACGACCCAGTCAGCTCTCAGACGAGACTTA 1977
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 1978 GGCCTCCAGCCCATGGACGACACCTTC 2007

RESULT 3
AL591390/c
LOCUS AL591390 186134 bp DNA linear ROD 17-NOV-2001
DEFINITION Mouse DNA sequence from clone RP23-355L10 on chromosome 11,
          complete sequence.
ACCESSION AL591390
VERSION AL591390.8 GI:17017767
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
  1 (Bases 1 to 186134)
  AUTHORS Phillimore,B.
  TITLE Direct Submission
  JOURNAL Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
          Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
          humquery@sanger.ac.uk clone requests: clonerequest@sanger.ac.uk
          On Nov 20, 2001 this sequence version replaced gi:16555512.
          During sequence assembly data is compared from overlapping clones.
          Where differences are found these are annotated as variations
          together with a note of the overlapping clone name; Note that the
          variation annotation may not be found in the sequence submission
          corresponding to the overlapping clone, as we submit sequences with
          only a small overlap as described above.
          This sequence was finished as follows unless otherwise noted: all
          regions were either double-stranded or sequenced with an alternate
          chemistry or covered by high quality data (i.e., phred quality >=
          30); an attempt was made to resolve all sequencing problems, such
          as compressions and repeats; all regions were covered by at least
          one plasmid subclone or more than one M13 subclone; and the
          assembly was confirmed by restriction digest. The following
          abbreviations are used to associate primary accession numbers given
          in the feature table with their source databases: Em:, EMBL; Sw:,
          SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; Information on the WORMPEP
          database can be found at
          http://www.sanger.ac.uk/projects/C-elegans/wormpep/
          from the RPCI-23 Mouse PAC Library
          constructed by the group of Pieter de Jong.
          For further details see http://www.chori.org/bacpac/home.htm
          VECTOR: pBACE3.6
          IMPORTANT: This sequence is not the entire insert of clone
          RP23-355L10 it may be shorter because we sequence overlapping
          sections only once, except for a short overlap.
          The true right end of clone RP23-355L10 is at 186134 in this
          sequence. The true right end of clone RP23-438D7 is at 2000 in this
          sequence.
FEATURES
    Location/Qualifiers
      1..186134
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10090"
      /chromosome="11"
      /clone="RP23-355L10"
      /clone_lib="RPCI-23"
      17991..18044
      /note="Sequence from uni-directional primer reads and dGTP
            big dye terminator reads only."
      47794..47812
      /note="Sequence from uni-directional primer reads and dGTP
            big dye terminator reads only."
BASE COUNT 46287 a 47900 c 47689 g 44258 t
ORIGIN

```









```

variation
902
/frequency="0.01"
/replace="a"
variation
975
/frequency="0.03"
/replace="t"
gene
1253..29789
/gene="ERBB2"
mRNA
join(1253..1421,8102..8253,9440..9653,10437..10571,
10932..11000,11205..11320,11459..11600,13047..13166,
13441..13567,16405..16478,16565..16655,16859..17058,
17420..17552,17634..17724,18439..18599,20908..20955,
24439..24577,24658..24780,25032..25130,25858..26043,
26181..26336,26459..26534,26839..26985,27694..27791,
27947..28135,28427..28679,28821..29789)
/gene="ERBB2"
product="v-erb-b2 erythroblastic leukemia viral oncogene
homolog 2, neuro/glioblastoma derived oncogene homolog
(avian)"
CDS
join(1349..1421,8102..8253,9440..9653,10437..10571,
10932..11000,11205..11320,11459..11600,13047..13166,
13441..13567,16405..16478,16565..16655,16859..17058,
17420..17552,17634..17724,18439..18599,20908..20955,
24439..24577,24658..24780,25032..25130,25858..26043,
26181..26336,26459..26534,26839..26985,27694..27791,
27947..28135,28427..28679,28821..29176)
/gene="ERBB2"
/codon_start=1
product="v-erb-b2 erythroblastic leukemia viral oncogene
homolog 2, neuro/glioblastoma derived oncogene homolog
(avian)"
/protein_id="AAO18082.1"
/db_xref="GI:27658000"
/translation="MELAALCEWGLLLALLPPGAASTOCTGTMKLRLPASPETHLD
MLRHLYQGVQVQGNLELYLPTNASLSFLQDIQVQGVVLIHNRQVPLQRIRV
RGTLQFDNYALAVLDGDPNLNTPVTGASPGGLRELQRLSLTELKGGVLIQRNPQ
LCYQDTILMKQNLALTIDNRSRACHGSPMCKGSCWGESSEDCQSILRT
VAGGACGKGLPTDCHEQCAAGCTGKSHSDCLALCFHNSGICELHCPALVYNT
DTFESMPREGRYTGASCTVACPNYLSTDVGCTPLHNSQVETAEQGTQCEK
SKPCARVCYGLMEHLREVRAVTSANIOEAGCKIFGLAFLPESFGDPASNTAPL
OPEOLQVETLEETGYLYISAWPDSLPLSVFONLOVIRGLHNGAYSLLTQGLGI
SWGLRSRLREGLGSLALIHNTILCFVHTVPWQDLFRNPQALLHTANRDECEGEG
LACHQLCARGHCWGPGTQVCNCSQFLRGQCEVEECRVLQGLPREYVNAHCLPCHPE
COPQNGSTCFGPEADQACAHYKDPFCVACRCPGVKPDLSYMWKFPDEEACQ
PCPINCHSDVDLDGCPAECASPLTSIIISAVVGLLVVGLVGVGILIKRROKI
RKYTWRRLLQETLEPLTPSGAMNQAMRLIKETLRKVVLGSGAFGVYKGIWI
PDGENVKIPVAIKVRENTSPKANKEILDEAVMAGVGSVYVRLGLICLSTVLVT
QLMPYGLLDHVRNRLGSDLLNWCWQIAKMSYLEDRVLRHDLAARNVLKSP
NWKITDFGLRLDIDETEHADGGKVPKWMALSIILRRFTHQSDVWSYGVTVWE
LMTFGAKPYDGIPIREIPDLLEKGERLPQPICTIDVYIMVWKMDSECRPREL
VSEFSRMARDPQRFVVIQNEIDLGASPLDSTFYRSLDDMDGDLVDAEYLVPOQGF
FCPDPAGAGGVHHRSSSTRSGGDLTLGLEPSEEPAPSPPLAPSEGAGSDVFDG
DLGMAAGLQSLPTHDPSPLQRYSEDTPVLPSETDGYVAPLCSPOPEYVNPQDVR
PQPSRREGPLPAAPAGATLERAKTSLFGKNGVVKDVFAGGAVENPEYLTPOGGA
PQHPPPAFSPAFDNLXYWDQDPPRGPSTFKGTPTAENPEYLGLDVPV"
1422..1712
/gene="ERBB2"
/note="Region not scanned for variation"
variation
1839
/gene="ERBB2"
/frequency="0.02"
/replace="c"
1887..6576
/gene="ERBB2"
/note="Region not scanned for variation"
misc_feature
2315..2536
/rpt_family="MIR"
/rpt_type="dispersed"
repeat_region
3217..3417
/rpt_family="Alu"
/rpt_type="dispersed"
repeat_region
3837..3974
/rpt_family="MIR"

```

```

repeat_region
4784..4858
/rpt_type="dispersed"
/rpt_family="MIR"
/rpt_type="dispersed"
5404..5693
/rpt_family="Alu"
/rpt_type="dispersed"
6074..6282
/rpt_family="MIR"
/rpt_type="dispersed"
6690
/gene="ERBB2"
/frequency="0.01"
/replace="t"
6952..6953
/gene="ERBB2"
/frequency="0.98"
/replace=""
6957
/gene="ERBB2"
/frequency="0.02"
/replace="g"
6960
/gene="ERBB2"
/frequency="0.02"
/replace="g"
7036
/gene="ERBB2"
/frequency="0.03"
/replace="t"
7146
/gene="ERBB2"
/frequency="0.01"
/replace="t"
7151
/gene="ERBB2"
/frequency="0.01"
/replace="t"
7825
/gene="ERBB2"
/frequency="0.06"
/replace="t"
7874
/gene="ERBB2"
/frequency="0.01"
/replace="a"
8127
/gene="ERBB2"
/frequency="0.01"
/replace="a"
8467..8473
/gene="ERBB2"
/frequency="0.96"
/replace=""
8574
/gene="ERBB2"
/frequency="0.02"
/replace="t"
8575..8707
/rpt_family="MIR"
/rpt_type="dispersed"
8836..9130
/rpt_family="Alu"
/rpt_type="dispersed"
8875..9119
/gene="ERBB2"
/note="Region not scanned for variation"
9136..9214
/rpt_family="MIR"
/rpt_type="dispersed"
9595
/gene="ERBB2"
/frequency="0.02"

```

```

/replace="g"
/702..9770
/rpt_family="tip100"
/rpt_type=dispersed
9766
/gene="ERBB2"
/frequency="0.01"
/replace="t"
/799..9946
/rpt_family="tip100"
/rpt_type=dispersed
9949..10151
/rpt_family="MIR"
/rpt_type=dispersed
10009
/gene="ERBB2"

```

```

Alignment Scores:
Pred. No.:      0.963      Length:      30837
Score:          85.00      Matches:      23
Percent Similarity: 46.00%      Conservative: 0
Best Local Similarity: 46.00%      Mismatches: 3
Query Match:      53.46%      Indels:      24
Db:              9          Gaps:      2

```

JUNC\_SEQ3\_SEQ4 (1-30) x AY208911 (1-30837)

```

QY  5  ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
      |||||
Db  27839 CCTGAGAGGTGGGAGGAGATGATGCATGATGCCAGCCCTCACGAGGCGTC 27898
QY  16 -----GlnAsnGluAspLeu 20
      |||||
Db  27899 ATGTGGCTGGGGGGGCCACCATCTGCTCTCTCTCTCCACAGATGAGGACTTG 27958
QY  21 GlyProAlaSerProLeuAspSerThrPhe 30
      |||||
Db  27959 GCGCCAGCCAGTCCCTTGGACAGCACCTTC 27988

```

```

RESULT 9
AC079199 AC079199 161815 bp DNA linear PRI 02-OCT-2002
LOCUS Homo sapiens chromosome 17, clone RP11-94L15, complete sequence.
DEFINITION AC079199
ACCESSION AC079199
VERSION AC079199.9 GI:23462913
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 161815)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP11-94L15
Unpublished
2 (bases 1 to 161815)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,N., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,

```

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (23-AUG-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 161815)  
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Direct Submission  
Submitted (08-SEP-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 161815)  
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
COMMENT

Submitted (02-OCT-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Oct 2, 2002 this sequence version replaced gi:22758800.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: LI0635  
Center clone name: 94\_L\_15  
-----

FEATURES  
source

Location/Qualifiers  
1..161815  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17"  
/clone="RP11-94L15"  
/clone\_lbb="RPC1-11 Human Male BAC"  
7..470  
/note="PCR product sequence only"

unsure

```

repeat_region complement(262..317)
unsure      /rpt_family="MIR"
repeat_region /note="PCR product sequence only"
repeat_region 1889..1986
repeat_region /rpt_family="AluSx"
repeat_region 5033..5179
repeat_region /rpt_family="MIR"
repeat_region 6113..6233
repeat_region /rpt_family="MIR"
repeat_region 6418..6460
repeat_region /rpt_family="L3"
repeat_region 7910..7930
repeat_region /rpt_family="AT-rich"
repeat_region 8422..8476
repeat_region /rpt_family="(CCTG)n"
repeat_region 8513..8659
repeat_region /rpt_family="MIR"
repeat_region complement(8917..9066)
repeat_region /rpt_family="MIR3"
repeat_region 12588..12697
repeat_region /rpt_family="G-rich"
repeat_region complement(13513..13824)
repeat_region /rpt_family="AluSp"
repeat_region complement(13836..13937)
repeat_region 18739..18939
repeat_region /rpt_family="MIR"
repeat_region 21075..21274
repeat_region /rpt_family="MIR"
repeat_region 21447..21714
repeat_region /rpt_family="MIR"
repeat_region complement(23385..23684)
repeat_region /rpt_family="AluSx"
repeat_region 24620..24819
repeat_region /rpt_family="MIR"
repeat_region 24991..25298
repeat_region /rpt_family="L2"
repeat_region complement(25299..25605)
repeat_region /rpt_family="AluJb"
repeat_region 25606..25686
repeat_region /rpt_family="L2"
repeat_region 25687..25838
repeat_region /rpt_family="MIR"
repeat_region 25839..25952
repeat_region /rpt_family="LTR54B"
repeat_region complement(25953..26241)
repeat_region /rpt_family="AluSp"
repeat_region 26242..26378
repeat_region /rpt_family="LTR54B"
repeat_region 26379..26695
repeat_region /rpt_family="AluY"
repeat_region 26696..26901
repeat_region /rpt_family="LTR54B"
repeat_region 26902..26985
repeat_region /rpt_family="MIR"
repeat_region 26986..27224
repeat_region /rpt_family="L2"
repeat_region 27243..27550
repeat_region /rpt_family="AluSg"
repeat_region 27695..28007
repeat_region /rpt_family="L2"
repeat_region 28290..28600
repeat_region /rpt_family="AluSx"
repeat_region 29462..29600
repeat_region /rpt_family="MIR"
repeat_region 29643..29967
repeat_region /rpt_family="AluY"
repeat_region complement(30700..31000)
repeat_region /rpt_family="AluSp"
repeat_region complement(31002..31248)
repeat_region /rpt_family="AluJb"
repeat_region 31521..31572

/rpt_family="L2"
31577..31796
/rpt_family="MIR"
32751..33072
/rpt_family="L2"
33181..33293
/rpt_family="L2"
34072..34376
/rpt_family="AluSp"
complement(34542..34833)
/rpt_family="AluJb"
complement(35218..35297)
/rpt_family="L2"
35320..35431
/rpt_family="AluSp"
35433..35500
/rpt_family="AluSp/q"
complement(35581..35702)
/rpt_family="L2"
35710..36008
/rpt_family="AluSc"
36009..36023

Alignment Scores:
Pred. No.:      5.18      Length:      161815
Score:          85.00      Matches:    23
Percent Similarity: 46.00%      Conservative: 0
Best Local Similarity: 46.00%      Mismatches: 3
Query Match:      53.46%      Indels:    24
DB:               9          Gaps:      2

JUNC_SEQ3_SEQ4 (1-30) x AC079199 (1-161815)
QY      5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
Db      71166 CCTGCAGAGGGTGGGAAGGAGAGATGATCCAGTATGCCAGGCCCTCAGGAAGCGTCG 71225
QY      16 -----GlnAsnGluAspLeu 20
Db      71226 ATGCTGGGCTGGGGAGGGGCCACCATCTTCCTCTCCATCCACAGATGAGGACTTG 71285
QY      21 GlyProAlaSerProLeuAspSerThrPhe 30
Db      71286 GGGCCAGCCAGTCCCTTGGACAGCACCTTC 71315

RESULT 10
AC040933
LOCUS      Homo sapiens chromosome 17, clone CTD-2019C10, complete sequence.
DEFINITION AC040933
ACCESSION AC040933
VERSION   AC040933.12 GI:23396303
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 168585)
Birren,B., Nusbaum,C. and Lander,E.
REFERENCE 2 (bases 1 to 168585)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
          Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
          Boguslavsky,L., Boukhgaiter,B., Brown,A., Burkett,G.,
          Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
          Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
          Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
          Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
          Grand-pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
          Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
          Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
          Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
          McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,

```

TITLE JOURNAL REFERENCE AUTHORS	Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severi, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.	FEATURES source	Location/Qualifiers 1. .168585 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="17" /map="17" /clone="CTD-2019C10" /clone_lib="CITD1 Human BAC" complement(1359. .1516) /rpt_family="MIR" complement(1538. .1754) /rpt_family="MIR" 2117. .2419 /rpt_family="AluSq" complement(2609. .2748) /rpt_family="MIR" 4091. .4278 /rpt_family="MIR" complement(4334. .4339) /note="<30 qual SNGL region" 4409. .4725 /rpt_family="AluJb" 4782. .4942 /rpt_family="L1ME1" 4986. .5046 /rpt_family="(TGAA)n" complement(5492. .5657) /rpt_family="MIR" 6438. .6626 /rpt_family="CT-rich" 8771. .8827 /rpt_family="GA-rich" 10677. .10875 /rpt_family="MIR3" complement(12314. .12419) /rpt_family="L2" 12443. .12588 /rpt_family="AluJo" 12589. .12885 /rpt_family="AluSp" 12886. .13057 /rpt_family="AluJo" /note="<30 qual SNGL region" complement(14521. .14553) complement(14571. .14592) /note="<30 qual SNGL region" complement(14609. .14614) /note="<30 qual SNGL region" complement(14624. .14629) /note="<30 qual SNGL region" complement(14639. .14666) /note="<30 qual SNGL region" 15796. .15849 /rpt_family="GA-rich" 16715. .17025 /rpt_family="AluSx" complement(17213. .17283) /rpt_family="MIR" complement(17647. .17651) /note="<30 qual SNGL region" 19655. .20204 /rpt_family="L2" 20443. .20649 /rpt_family="MIR" 20863. .20928 /rpt_family="AluJo" 21295. .21301 /note="<30 qual SNGL region" 21587. .21929 /rpt_family="L2" complement(22041. .22147) /rpt_family="MIR"
	Direct Submission		
	Submitted (11-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 3 (bases 1 to 168585)		
TITLE JOURNAL REFERENCE AUTHORS	Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	FEATURES source	Location/Qualifiers 1. .168585 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="17" /map="17" /clone="CTD-2019C10" /clone_lib="CITD1 Human BAC" complement(1359. .1516) /rpt_family="MIR" complement(1538. .1754) /rpt_family="MIR" 2117. .2419 /rpt_family="AluSq" complement(2609. .2748) /rpt_family="MIR" 4091. .4278 /rpt_family="MIR" complement(4334. .4339) /note="<30 qual SNGL region" 4409. .4725 /rpt_family="AluJb" 4782. .4942 /rpt_family="L1ME1" 4986. .5046 /rpt_family="(TGAA)n" complement(5492. .5657) /rpt_family="MIR" 6438. .6626 /rpt_family="CT-rich" 8771. .8827 /rpt_family="GA-rich" 10677. .10875 /rpt_family="MIR3" complement(12314. .12419) /rpt_family="L2" 12443. .12588 /rpt_family="AluJo" 12589. .12885 /rpt_family="AluSp" 12886. .13057 /rpt_family="AluJo" /note="<30 qual SNGL region" complement(14521. .14553) complement(14571. .14592) /note="<30 qual SNGL region" complement(14609. .14614) /note="<30 qual SNGL region" complement(14624. .14629) /note="<30 qual SNGL region" complement(14639. .14666) /note="<30 qual SNGL region" 15796. .15849 /rpt_family="GA-rich" 16715. .17025 /rpt_family="AluSx" complement(17213. .17283) /rpt_family="MIR" complement(17647. .17651) /note="<30 qual SNGL region" 19655. .20204 /rpt_family="L2" 20443. .20649 /rpt_family="MIR" 20863. .20928 /rpt_family="AluJo" 21295. .21301 /note="<30 qual SNGL region" 21587. .21929 /rpt_family="L2" complement(22041. .22147) /rpt_family="MIR"
	Direct Submission		
	Submitted (01-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 168585)		
TITLE JOURNAL REFERENCE AUTHORS	Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.	FEATURES source	Location/Qualifiers 1. .168585 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="17" /map="17" /clone="CTD-2019C10" /clone_lib="CITD1 Human BAC" complement(1359. .1516) /rpt_family="MIR" complement(1538. .1754) /rpt_family="MIR" 2117. .2419 /rpt_family="AluSq" complement(2609. .2748) /rpt_family="MIR" 4091. .4278 /rpt_family="MIR" complement(4334. .4339) /note="<30 qual SNGL region" 4409. .4725 /rpt_family="AluJb" 4782. .4942 /rpt_family="L1ME1" 4986. .5046 /rpt_family="(TGAA)n" complement(5492. .5657) /rpt_family="MIR" 6438. .6626 /rpt_family="CT-rich" 8771. .8827 /rpt_family="GA-rich" 10677. .10875 /rpt_family="MIR3" complement(12314. .12419) /rpt_family="L2" 12443. .12588 /rpt_family="AluJo" 12589. .12885 /rpt_family="AluSp" 12886. .13057 /rpt_family="AluJo" /note="<30 qual SNGL region" complement(14521. .14553) complement(14571. .14592) /note="<30 qual SNGL region" complement(14609. .14614) /note="<30 qual SNGL region" complement(14624. .14629) /note="<30 qual SNGL region" complement(14639. .14666) /note="<30 qual SNGL region" 15796. .15849 /rpt_family="GA-rich" 16715. .17025 /rpt_family="AluSx" complement(17213. .17283) /rpt_family="MIR" complement(17647. .17651) /note="<30 qual SNGL region" 19655. .20204 /rpt_family="L2" 20443. .20649 /rpt_family="MIR" 20863. .20928 /rpt_family="AluJo" 21295. .21301 /note="<30 qual SNGL region" 21587. .21929 /rpt_family="L2" complement(22041. .22147) /rpt_family="MIR"
	Direct Submission		
	Submitted (01-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Oct 1, 2002 this sequence version replaced gi:22597589. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) <a href="http://ftp.genome.washington.edu/RM/RepeatMasker.html">http://ftp.genome.washington.edu/RM/RepeatMasker.html</a>		
TITLE JOURNAL COMMENT	Center code: WIBR Web site: <a href="http://www.seq.wi.mit.edu">http://www.seq.wi.mit.edu</a> Contact: <a href="mailto:sequence_submissions@genome.wi.mit.edu">sequence_submissions@genome.wi.mit.edu</a> Project Information Center project name: L8894 Center clone name: 2019_C10	FEATURES source	Location/Qualifiers 1. .168585 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="17" /map="17" /clone="CTD-2019C10" /clone_lib="CITD1 Human BAC" complement(1359. .1516) /rpt_family="MIR" complement(1538. .1754) /rpt_family="MIR" 2117. .2419 /rpt_family="AluSq" complement(2609. .2748) /rpt_family="MIR" 4091. .4278 /rpt_family="MIR" complement(4334. .4339) /note="<30 qual SNGL region" 4409. .4725 /rpt_family="AluJb" 4782. .4942 /rpt_family="L1ME1" 4986. .5046 /rpt_family="(TGAA)n" complement(5492. .5657) /rpt_family="MIR" 6438. .6626 /rpt_family="CT-rich" 8771. .8827 /rpt_family="GA-rich" 10677. .10875 /rpt_family="MIR3" complement(12314. .12419) /rpt_family="L2" 12443. .12588 /rpt_family="AluJo" 12589. .12885 /rpt_family="AluSp" 12886. .13057 /rpt_family="AluJo" /note="<30 qual SNGL region" complement(14521. .14553) complement(14571. .14592) /note="<30 qual SNGL region" complement(14609. .14614) /note="<30 qual SNGL region" complement(14624. .14629) /note="<30 qual SNGL region" complement(14639. .14666) /note="<30 qual SNGL region" 15796. .15849 /rpt_family="GA-rich" 16715. .17025 /rpt_family="AluSx" complement(17213. .17283) /rpt_family="MIR" complement(17647. .17651) /note="<30 qual SNGL region" 19655. .20204 /rpt_family="L2" 20443. .20649 /rpt_family="MIR" 20863. .20928 /rpt_family="AluJo" 21295. .21301 /note="<30 qual SNGL region" 21587. .21929 /rpt_family="L2" complement(22041. .22147) /rpt_family="MIR"
	Center code: WIBR Web site: <a href="http://www.seq.wi.mit.edu">http://www.seq.wi.mit.edu</a> Contact: <a href="mailto:sequence_submissions@genome.wi.mit.edu">sequence_submissions@genome.wi.mit.edu</a> Project Information Center project name: L8894 Center clone name: 2019_C10		
	Center code: WIBR Web site: <a href="http://www.seq.wi.mit.edu">http://www.seq.wi.mit.edu</a> Contact: <a href="mailto:sequence_submissions@genome.wi.mit.edu">sequence_submissions@genome.wi.mit.edu</a> Project Information Center project name: L8894 Center clone name: 2019_C10		

Only the last 167.6 kb of this clone are being submitted.  
The remainder overlaps accession number AC097491 [WICGR project L11967].

```

repeat_region 22878..22922
/rpt_family="AluYb8"
repeat_region 23134..23269
/rpt_family="FAM"
repeat_region complement(23439..23573)
/rpt_family="MIR"
repeat_region complement(23801..24002)
/rpt_family="L2"
repeat_region complement(24332..25035)
/rpt_family="L2"
repeat_region complement(25598..25820)
/rpt_family="MIR"
repeat_region complement(26182..26352)
/rpt_family="MIR"
repeat_region 26378..26669
/rpt_family="AluJo"
repeat_region complement(27092..27193)
/rpt_family="MIR"
repeat_region complement(27266..27543)
/rpt_family="AluX"
repeat_region complement(27656..27682)
/rpt_family="MIR"
repeat_region complement(27683..27991)
/rpt_family="AluX"
repeat_region complement(27999..28048)
/rpt_family="LMB4"
repeat_region complement(28049..28353)
/rpt_family="AluJb"

Alignment Scores:
Pred. No.: 5.4 Length: 168585
Score: 85.00 Matches: 23
Percent Similarity: 46.00% Conservative: 0
Best Local Similarity: 46.00% Mismatches: 3
Query Match: 53.46% Indels: 24
DB: 9 Gaps: 2

JUNC_SEQ3_SEQ4 (1-30) x AC040933 (1-168585)
QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
Db 108481 CCTGCAGAGGGTGGGAAGGAGAGATGATGCCAGGCCCTCAGCGAAGGCTGC 108540
QY 16 -----GlnAsnGluAspLeu 20
Db 108541 ATGCTGGGCTGGGAGGGCCACCATCTGCTCTCTCTCCCTCCACAGATCAGGACTTG 108600
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 108601 GGCCAGCAGCTCCCTTGGACAGCACCTTC 108630

RESULT 11
AC142197
LOCUS AC142197 198008 bp DNA linear HTG 25-MAR-2003
DEFINITION Homo sapiens chromosome 16 clone RP11-1054B13, WORKING DRAFT
SEQUENCE, 8 unordered pieces.
ACCESSION AC142197
VERSION AC142197.1 GI:29171373
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 198008)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 198008)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-2003) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center

```

```

Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1699277
Center clone name: RPCI-11_1054B13
-----
Summary Statistics
Consensus quality: 193482 bases at least Q40
Consensus quality: 194189 bases at least Q30
Consensus quality: 194769 bases at least Q20
Estimated insert size: 175000; agarose-fp estimation
Estimated insert size: 197308; sum-of-contigs estimation
Quality coverage: 10.95 in Q20 bases; agarose-fp estimation
Quality coverage: 9.72 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 1185: contig of 1185 bp in length
* 1186 1285: gap of unknown length
* 1286 5050: contig of 3765 bp in length
* 5051 5150: gap of unknown length
* 5151 10780: contig of 5630 bp in length
* 10781 10880: gap of unknown length
* 10881 24702: contig of 13822 bp in length
* 24703 24802: gap of unknown length
* 24803 56555: contig of 31853 bp in length
* 56556 56755: gap of unknown length
* 56756 100385: contig of 43630 bp in length
* 100386 100485: gap of unknown length
* 100486 144983: contig of 44498 bp in length
* 144984 145083: gap of unknown length
* 145084 198008: contig of 52925 bp in length.
FEATURES
Location/Qualifiers
1..198008
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-1054B13"
/clone_lib="RPCI human BAC library 11"
BASE COUNT 53974 a 44750 c 45309 g 53270 t 705 others
ORIGIN
Alignment Scores:
Pred. No.: 6.35 Length: 198008
Score: 85.00 Matches: 23
Percent Similarity: 46.00% Conservative: 0
Best Local Similarity: 46.00% Mismatches: 3
Query Match: 53.46% Indels: 24
DB: 2 Gaps: 2

JUNC_SEQ3_SEQ4 (1-30) x AC142197 (1-198008)
QY 5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
Db 191973 CCTGCAGAGGGTGGGAAGGAGAGATGATGCCAGGCCCTCAGCGAAGGCTGC 192032
QY 16 -----GlnAsnGluAspLeu 20
Db 192033 ATGCTGGGCTGGGAGGGGCCACCATCTCTCTCTCCCTCCACAGATCAGGACTTG 192092
QY 21 GlyProAlaSerProLeuAspSerThrPhe 30
Db 192093 GGCCAGCAGCTCCCTTGGACAGCACCTTC 192122
RESULT 12

```

```
AR082744      AR082744      2385 bp      DNA      linear      PAT 01-SEP-2000
LOCUS          Sequence 3 from patent US 5976546.
ACCESSION      AR082744
VERSION        AR082744.1  GI:10009534
KEYWORDS       Unknown.
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE      1 (bases 1 to 2385)
AUTHORS        Laus,R., Ruegg,C.Landon. and Wu,H.
TITLE          Immunostimulatory compositions
JOURNAL        Patent: US 5976546-A 3 02-NOV-1999;
FEATURES       Location/Qualifiers
               source
               1..2385
               /organism="unknown"
BASE COUNT    488 a 780 c 677 g 440 t
ORIGIN
Alignment Scores:
Pred. No.:      0.098      Length:      2385
Score:          84.00      Matches:    21
Percent Similarity: 70.00%      Conservative: 0
Best Local Similarity: 70.00%      Mismatches: 5
Query Match:    52.83%      Indels:    4
DB:             6      Gaps:      2

JUNC_SEQ3_SEQ4 (1-30) x AR082744 (1-2385)
Qy 1 AsplyGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1925 GACAAGGGCTGCCCGCGAGAGAGAGCCAGCCCTCTGAGCTCCCTCGAG----- 1975

Qy 21 GlyProAla---SerProLeuAspSerThr 29
Db 1976 GCACCGCGCGCTCGCCCGAGCCAGCCAGCACA 2005

RESULT 13
AR099963
LOCUS          Sequence 3 from patent US 6080409.
ACCESSION      AR099963
VERSION        AR099963.1  GI:12810411
KEYWORDS       Unknown.
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE      1 (bases 1 to 2385)
AUTHORS        Laus,R., Ruegg,C.Landon. and Wu,H.
TITLE          Immunostimulatory method
JOURNAL        Patent: US 6080409-A 3 27-JUN-2000;
FEATURES       Location/Qualifiers
               source
               1..2385
               /organism="unknown"
BASE COUNT    488 a 780 c 677 g 440 t
ORIGIN
Alignment Scores:
Pred. No.:      0.098      Length:      2385
Score:          84.00      Matches:    21
Percent Similarity: 70.00%      Conservative: 0
Best Local Similarity: 70.00%      Mismatches: 5
Query Match:    52.83%      Indels:    4
DB:             6      Gaps:      2

JUNC_SEQ3_SEQ4 (1-30) x AR099963 (1-2385)
Qy 1 AsplyGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1925 GACAAGGGCTGCCCGCGAGAGAGCCAGCCCTCTGAGCTCCCTCGAG----- 1975

Qy 21 GlyProAla---SerProLeuAspSerThr 29
Db 1976 GCACCGCGCGCTCGCCCGAGCCAGCCAGCACA 2005

RESULT 14
AR143949
LOCUS          Sequence 3 from patent US 6210662.
ACCESSION      AR143949
VERSION        AR143949.1  GI:15105816
KEYWORDS       Unknown.
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE      1 (bases 1 to 2385)
AUTHORS        Laus,R., Ruegg,C.Landon. and Wu,H.
TITLE          Immunostimulatory composition
JOURNAL        Patent: US 6210662-A 3 03-APR-2001;
FEATURES       Location/Qualifiers
               source
               1..2385
               /organism="unknown"
BASE COUNT    488 a 780 c 677 g 440 t
ORIGIN
Alignment Scores:
Pred. No.:      0.098      Length:      2385
Score:          84.00      Matches:    21
Percent Similarity: 70.00%      Conservative: 0
Best Local Similarity: 70.00%      Mismatches: 5
Query Match:    52.83%      Indels:    4
DB:             6      Gaps:      2

JUNC_SEQ3_SEQ4 (1-30) x AR143949 (1-2385)
Qy 1 AsplyGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1925 GACAAGGGCTGCCCGCGAGAGAGCCAGCCCTCTGAGCTCCCTCGAG----- 1975

Qy 21 GlyProAla---SerProLeuAspSerThr 29
Db 1976 GCACCGCGCGCTCGCCCGAGCCAGCCAGCACA 2005

RESULT 15
AX022984
LOCUS          Sequence 4 from Patent EP0919622.
ACCESSION      AX022984
VERSION        AX022984.1  GI:10046479
KEYWORDS       unidentified
SOURCE         unidentified
ORGANISM       unclassified.
REFERENCE      1
AUTHORS        Springer,C.J. and Marais,R.
TITLE          Surface expression of enzyme in gene directed prodrug therapy
JOURNAL        Patent: EP 0919622-A 4 02-JUN-1999;
FEATURES       Location/Qualifiers
               source
               1..153
               /organism="unidentified"
               /mol_type="genomic DNA"
               /db_xref="taxon:32644"
BASE COUNT    27 a 43 c 51 g 32 t
ORIGIN
Alignment Scores:
Pred. No.:      0.021      Length:      153
Score:          80.00      Matches:    15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    50.31%      Indels:    0
DB:             6      Gaps:      0

JUNC_SEQ3_SEQ4 (1-30) x AX022984 (1-153)
```

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
Db 10 GACAAGGCTGCCCGCGGAGCAGAGAGCCAGCCCTCTGAGGTCC 54

Search completed: October 16, 2003, 10:59:54  
Job time : 503.208 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.  
OM protein - nucleic search, using frame\_plus.p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 234.794 Seconds  
(without alignments)  
3105.423 Million cell updates/sec

Title: JUNC\_SEQ3\_SEQ2RES991\_  
Perfect score: 160  
Sequence: 1 DRGCPAQRASPLTSQNDLGPSPMDSTF 30  
Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0  
Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DRV=slp  
-Q/cgn2\_1/USPTO.spool.p/HOLLERAN480/runat\_15102003\_131913\_20548/app\_query.fasta\_1.4685  
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 -CGN\_1\_1\_12645 -runat\_15102003\_131913\_20548 -NCPU=6 -ICPU=3  
-NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estnu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_htc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*

29: gb\_gss2.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	91.5	57.2	618	28	AZ699313	AZ699313 RPCI-23-2
2	83.5	52.2	673	12	BI157930	BI157930 602920909
3	80	50.0	378	10	BF869854	BF869854 IL3-ET011
4	80	50.0	411	2	HS0069572	Bx479682 Homo sapi
5	80	50.0	422	9	AW606421	AW606421 RC3-HT037
6	80	50.0	422	10	BF554233	BF554233 UI-R-C0-h
7	80	50.0	433	10	BF762317	BF762317 IL2-CS004
8	80	50.0	439	10	BE157169	BE157169 RC3-HT037
9	80	50.0	450	12	BG988031	BG988031 MR2-HT116
10	80	50.0	466	12	BM720186	BM720186 UI-E-E00-
11	80	50.0	524	13	BQ828479	BQ828479 LL6in1002
12	80	50.0	527	14	CB719751	CB719751 AMGNNUC:N
13	80	50.0	546	12	BI194790	BI194790 602948150
14	80	50.0	547	9	AA443351	AA443351 zw85b05.r
15	80	50.0	561	12	BM720098	BM720098 UI-E-E00-
16	80	50.0	568	13	BQ831943	BQ831943 LL6in1122
17	80	50.0	577	13	BQ829915	BQ829915 LL6in2239
18	80	50.0	592	10	BE382065	BE382065 601272225
19	80	50.0	607	12	BM083444	BM083444 Imageqc_6
20	80	50.0	613	9	AW410534	AW410534 fh06h06.x
21	80	50.0	629	10	BE374596	BE374596 601225384
22	80	50.0	666	13	BU671195	BU671195 NISC_lr06
23	80	50.0	667	12	BM048624	BM048624 603623984
24	80	50.0	674	12	BM758528	BM758528 K-EST0038
25	80	50.0	715	13	BU052728	BU052728 UI-M-E00-
26	80	50.0	715	13	BI648082	BI648082 603279290
27	80	50.0	726	12	BE15590	BE15590 601278973
28	80	50.0	741	10	BE15590	BE15590 601278973
29	80	50.0	770	12	BI645946	BI645946 6032921306
30	80	50.0	773	12	BI157032	BI157032 602921306
31	80	50.0	778	12	BG863431	BG863431 60297093
32	80	50.0	785	12	BI688650	BI688650 603311262
33	80	50.0	787	12	BI156730	BI156730 602922150
34	80	50.0	844	12	BI158626	BI158626 602922395
35	80	50.0	881	9	AI906106	AI906106 RC-BT105-
36	80	50.0	881	12	BI645770	BI645770 603275203
37	80	50.0	886	13	BQ882315	BQ882315 AGENCOURT
38	80	50.0	887	14	CA980253	CA980253 AGENCOURT
39	80	50.0	893	14	CA487631	CA487631 AGENCOURT
40	80	50.0	905	13	BU541491	BU541491 AGENCOURT
41	80	50.0	909	10	BG289457	BG289457 602381495
42	80	50.0	922	14	CA453990	CA453990 AGENCOURT
43	80	50.0	944	14	CD558559	CD558559 AGENCOURT
44	80	50.0	984	13	BU523863	BU523863 AGENCOURT
45	80	50.0	1009	13	BQ647343	BQ647343 AGENCOURT

ALIGNMENTS

RESULT 1  
AZ699313 618 bp DNA linear GSS 24-JAN-2001  
LOCUS RPCI-23-240N16.TJ RPCI-23 Mus musculus genomic clone RPCI-23-240N16  
DEFINITION , genomic survey sequence.  
ACCESSION AZ699313  
VERSION AZ699313.1 GI:12419623  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 618)

**AUTHORS** Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akintret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P., and Fraser, C. M.  
**TITLE** Mouse BAC End Sequences from Library RPCI-23  
**JOURNAL** Unpublished  
**COMMENT** Other\_GSSS: RPCI-23-240N16.TV  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org  
 Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pjejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: [http://www.tigr.org/tldb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html)  
 Plate: 240 row: N column: 16  
 Seq primer: SP6  
 Class: BAC ends.

**FEATURES** Location/Qualifiers  
 1..618  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-23-240N16"  
 /sex="Female"  
 /lab\_host="DH10B"  
 /clone\_lib="RPCI-23"  
 /note="Organ: Kidney/Brain; Vector: pBACE3.6; Site.1: EcoRI; Site.2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methyase. Size selected DNA was cloned into the pBACE3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."  
 BASE COUNT 110 a 214 c 151 g 143 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.057 Length: 618  
 Score: 91.50 Matches: 20  
 Percent Similarity: 72.41% Conservative: 1  
 Best Local Similarity: 68.97% Mismatches: 5  
 Query Match: 57.19% Indels: 3  
 DB: 28 Gaps: 1

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x AZ699313 (1-618)

QY 2 LysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGly 21  
 Db 165 GAGGGCTGT-----CACCTCGTGTCTCCCTTCTCTCTAGACGAGGACTTAGGC 215  
 QY 22 ProSerProMetAspSerThrPhe 30  
 Db 216 CCTCCAGCCCATGGACGACACCTTC 242

**RESULT 2**  
 B1157930  
 LOCUS B1157930 673 bp mRNA linear EST 05-JUL-2001  
 DEFINITION 602920909F1 NCI\_CGAP\_Mam3 Mus musculus cDNA clone IMAGE:5061027 5', mRNA sequence.  
 ACCESSION B1157930  
 VERSION B1157930.1 GI:14617931  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus.  
 REFERENCE 1 (bases 1 to 673)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

**TITLE** JOURNAL  
**COMMENT**

National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LLAM1165 row: b column: 04  
 High quality sequence stop: 664.

**FEATURES** Location/Qualifiers  
 1..673  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="129,C57BL/6J,FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5061027"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="10 months"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP\_Mam3"  
 /note="Organ: mammary; Vector: pCMV-SPORT5; Site.1: SalI; Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."  
 BASE COUNT 146 a 198 c 181 g 148 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.668 Length: 673  
 Score: 83.50 Matches: 20  
 Percent Similarity: 73.33% Conservative: 2  
 Best Local Similarity: 66.67% Mismatches: 4  
 Query Match: 52.19% Indels: 4  
 DB: 12 Gaps: 1

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x B1157930 (1-673)

QY 2 LysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGly 21  
 Db 473 GAGGGCTGT-----CACCTCGTGTCTCCCTTCTCTCTAGAACGAGGACTTAGG 523  
 QY 21 yProSerProMetAspSerThrPhe 30  
 Db 524 CCCCTCCAGCCCATGGACGACACCTTC 551

**RESULT 3**  
 B1157930  
 LOCUS B1157930 378 bp mRNA linear EST 17-JAN-2001  
 DEFINITION IL3-ET0114-251000-317-B07 ET0114 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION B1157930  
 VERSION B1157930.1 GI:12259984  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 378)  
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsumura, A., Baia, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.  
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663



```

/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0371"
/Note="Organ: head neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT      77 a 126 c 135 g 84 t
ORIGIN

Alignment Scores:
Pred. No.:      1.04      Length:      422
Score:          80.00     Matches:      21
Percent Similarity: 46.00%   Conservative: 2
Best Local Similarity: 42.00% Mismatches: 3
Query Match:    50.00%   Indels:      24
DB:              9      Gaps:        2

JUNC_SEQ3_SEQ2RES991_ (1-30) x AW606421 (1-422)
.
QY      5 ProAlaGlu-----GlnArgAlaSerProLeuThrSer----- 15
      |||||
Db      256 CCTGCAGAGGGTGGGAAGGAGAGATGATGCCAGTATGCCAGGCCCTCAGCGAAGGGCTGC 315
QY      16 -----GlnAsnGluAspLeu 20
      |||||
Db      316 ATGCTGGGCTGGGAGGGGCCACATCGCTCTCTCTCCATCCACAGATGAGGACTTG 375

QY      21 GlyProSerProMetAspSerThrPhe 30
      |||||
Db      376 GCGCCAGCCAGTCCCTTGGACAGCACCTTC 405

RESULT 6
BF554233
LOCUS      422 bp      mRNA      linear      EST 12-DEC-2000
DEFINITION      UI-R-C0-hp-c-01-0-UI-r1 UI-R-C0 Rattus norvegicus cDNA clone
ACCESSION      BF554233
VERSION      BF554233.1 GI:11663963
KEYWORDS
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
REFERENCE      1 (bases 1 to 422)
AUTHORS      Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      9704477
PUBMED      8889548
COMMENT      Contact: Soares, MB
      Coordinated Laboratory for Computational Genomics
      University of Iowa
      375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
      Tel: 319 335 8250
      Fax: 319 335 9565
      Email: bento-soares@uiowa.edu
      cDNA Library Preparation: M.B. Soares Lab Clone distribution:
      clones will be available through Research Genetics (www.resgen.com)
      This clone is also available through the I.M.A.G.E. Consortium at
      LLNL (infoimage.llnl.gov). IMAGE ID= 1781731
      Seq primer: M13 Forward.
      Location/Qualifiers
      1. .422
      /organism="Rattus norvegicus"
      /mol_type="mRNA"
      /strain="Sprague-Dawley"
      /db_xref="taxon:10116"

```

```

/clone="UI-R-C0-hp-c-01-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-C0"
/Note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-C0 library is a subtracted library derived from the UI-R-A1 and UI-R-E1 libraries. The UI-R-A1 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, and muscle. The UI-R-E1 library consisted of a mixture of individually tagged normalized libraries constructed from 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C0) was constructed as follows: PCR amplified cDNA inserts from a pool of UI-R-A1 and UI-R-E1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the pooled UI-R-A1 and UI-R-E1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C0 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)."
BASE COUNT      90 a 115 c 125 g 92 t
ORIGIN

Alignment Scores:
Pred. No.:      1.04      Length:      422
Score:          80.00     Matches:      15
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    50.00%   Indels:      0
DB:              10      Gaps:        0

JUNC_SEQ3_SEQ2RES991_ (1-30) x BF554233 (1-422)
QY      16 GlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
      |||||
Db      131 CAGACGAGGACTTGGCCCATCCAGCCCATGGACAGTACCTTC 175

RESULT 7
BF762317/c
LOCUS      433 bp      mRNA      linear      EST 12-JAN-2001
DEFINITION      IL2-CS0049-181000-202-A06 CS0049 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BF762317
VERSION      BF762317.1 GI:12110217
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS      1 (bases 1 to 433)
      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
      Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
PUBMED      10737800
COMMENT      Contact: Simpson A.J.G.
      Laboratory of Cancer Genetics
      Ludwig Institute for Cancer Research
      Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

```



Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-HT1160-110101-005-a06&t3=2001-01-11&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 39  
 High quality sequence stop: 325.

# FEATURES

Location/Qualifiers  
 1..450  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HT1160"  
 /note="Organ: head\_neck; Vector: puc18; Site: SmaI; Site2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196.716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
 BASE COUNT 84 a 127 c 150 g 89 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.13 Length: 450  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.00% Indels: 0  
 DB: 12 Gaps: 0  
 JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x BG988031 (1-450)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
 |||||  
 Db 259 GACAAGGGCTGCCCTGCCGAGCAGAGCCGCCCTCTGACGTC 303

RESULT 10  
 BM720186  
 LOCUS  
 DEFINITION  
 UI-E-E00-ahz-c-08-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone  
 UI-E-E00-ahz-c-08-0-UI 5', mRNA sequence.  
 BM720186  
 VERSION  
 BM720186.1 GI:19039265  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 466)  
 Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 97044477.  
 8889548  
 Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 Reverse.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 COMMENT

Location/Qualifiers  
 1..466  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-E00-ahz-c-08-0-UI"  
 /tissue\_type="fetal eye"  
 /dev\_stage="fetal"  
 /lab\_host="DHI0B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-E00"  
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site1: Ecor I; Site2: Not I; UI-E-E00 is a cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is CCGGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."  
 BASE COUNT 105 a 114 c 151 g 93 t 3 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.18 Length: 466  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.00% Indels: 0  
 DB: 12 Gaps: 0  
 JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x BM720186 (1-466)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
 |||||  
 Db 18 GACAAGGGCTGCCCTGCCGAGCAGAGCCGCCCTCTGACGTC 62

RESULT 11  
 BQ828479  
 LOCUS  
 DEFINITION  
 LL6in10025T7 AFT024-subtracted library Mus musculus cDNA 5' similar  
 to Neu/ERBB-2 p185, mRNA sequence.  
 BQ828479  
 ACCESSION  
 BQ828479  
 KEYWORDS  
 EST.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 524)  
 Hackney,J.A., Charbord,P., Brunk,B.P., Stoeckert,C.J., Lemischka  
 ,I.R. and Moore,K.A.  
 A molecular profile of a hematopoietic stem cell niche  
 Proc. Natl. Acad. Sci. U.S.A. 99 (20), 13061-13066 (2002)  
 12226475  
 PUBMED  
 COMMENT  
 Contact: Moore, Kateri A.  
 Department of Molecular Biology  
 Princeton University  
 217 Lewis Thomas Laboratory, Washington Road, Princeton, NJ 08544,  
 USA  
 Tel: 609 258 0605  
 Fax: 609 258 2759  
 Email: kmoore@molbio.princeton.edu  
 These ESTs are derived from a subtracted cDNA library enriched for  
 gene products expressed by a hematopoietic stem cell-supporting  
 stromal cell line, AFT024.  
 Seq primer: M13Reverse or T7.

FEATURES  
source

Location/Qualifiers  
1. .524  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57Bl/6J"  
/db\_xref="taxon:10090"  
/tissue\_type="Fetal Liver"  
/cell\_type="Stromal cell"  
/cell\_line="AFT024"  
/dev\_stage="Embryonic day 14-14.5"  
/lab\_host="DH108"  
/clone\_lib="AFT024-subtracted library"  
/note="Organ: Fetal Liver; Vector: Sport 1; Site\_1: Sal I; Site\_2: Not I; Two directionally cloned cDNA libraries were made from a hematopoietic stem cell-supporting stromal cell line (AFT024) and from a non-supporting stromal cell line (2018). Subtractive hybridization was performed by hybridization of the target, AFT024, single stranded cDNA library in pSport1 to biotinylated RNA transcribed from the driver, 2018 cDNA library in pSport2 with inserts cloned in the complementary orientation. The AFT024-subtracted library contains 4.2x105 clones and is depleted of common housekeeping gene products eg. beta-actin and enriched for transcripts specific to AFT024. For detailed protocols and additional information please see our website at <http://stromalcell.princeton.edu>."

BASE COUNT 110 a 149 c 157 g 107 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.: 1.37 Length: 524  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.00% Indels: 0  
DB: 13 Gaps: 0

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x BQ828479 (1-524)

QY 16 GlnAsGluAspLeuGlyProSerSerProMetAspSerThrphe 30  
|||||  
Db 279 CAGACGAGGACTAGGCCCTCCAGCCCATGGACGACACCTTC 323

## RESULT 12

LOCUS CB719751 527 bp mRNA linear EST 10-APR-2003  
DEFINITION AMGNNUC:NRD01-00186-c7-A nrdg1 (10855) Rattus norvegicus cDNA clone nrdg1-00186-c7 5', mRNA sequence.

ACCESSION CB719751.1 GI:29776893

VERSION CB719751.1

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 527)

Angen EST Program.

Angen Rat EST Program

Unpublished

JOURNAL Contact: Dan Fitzpatrick

COMMENT Angen, Inc

One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA

Tel: 805 447-4881

Plate: 00186 row: c column: 7.

Location/Qualifiers

1. .527

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

/clone="nrdg1-00186-c7"

/tissue\_type="Dorsal Root Ganglia"

FEATURES  
source

BASE COUNT 110 a 147 c 183 g 106 t  
ORIGIN

## Alignment Scores:

Pred. No.: 1.44 Length: 546  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0

/clone\_lib="nrdg1 (10855)"  
/note="Vector: pSPORT1; Site\_1: SalI; Site\_2: NotI; rat dorsal root ganglia"  
BASE COUNT 105 a 159 c 156 g 107 t  
ORIGIN

## Alignment Scores:

Pred. No.: 1.38 Length: 527  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x CB719751 (1-527)

QY 16 GlnAsGluAspLeuGlyProSerSerProMetAspSerThrphe 30  
|||||

Db 96 CAGACGAGGACTGGGCCCATCCAGCCCATGGACGACGACCTTC 140

## RESULT 13

LOCUS BI194790 546 bp mRNA linear EST 10-JUL-2001  
DEFINITION 602948150F1 NIH\_MGC\_42 Homo sapiens cDNA clone IMAGE:5091396 5', mRNA sequence.

ACCESSION BI194790

VERSION BI194790.1 GI:14649810

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 546)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCMI855 row: c column: 13

High quality sequence start: 10

High quality sequence stop: 495.

Location/Qualifiers

1. .546

source

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5091396"

/tissue\_type="epithelioid carcinoma cell line"

/lab\_host="DH108 (phage-resistant)"

/clone\_lib="NIH\_MGC\_42"

/note="Organ: pancreas; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling

Hong in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library. "

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.00% Indels: 0  
 DB: 12 Gaps: 0

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x BI194790 (1-546)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
 |||||  
 Db 223 GACAAGGGCTGCCCGCGAGAGAGAGCCACCCCTCTGACGTCC 267

## RESULT 14

AA443351

LOCUS

DEFINITION zW85b05.r1 Soares total.fetus.Nb2HF8.9w Homo sapiens cDNA clone  
 IMAGE:783729 5' similar to gb:M11730 ERBB-2 RECEPTOR  
 PROTEIN-TYROSINE KINASE PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION AA443351

VERSION AA443351.1 GI:2156026

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 547)  
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,  
 Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie,  
 T., Waterston, R., and Wilson, R.

TITLE WashU-Merck EST Project 1997

JOURNAL Unpublished

COMMENT

Contact: Wilison RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 360.

## FEATURES

source

1..547

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:5981347"

/db\_xref="taxon:9606"

/clone="IMAGE:783729"

/dev\_stage="8-9 weeks"

/lab\_host="DH10B"

/clone\_lib="Soares total.fetus.Nb2HF8.9w"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
 polylinker; Site1: Not I; Site2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from pooled 8-9 week  
 (total) fetus material with a Not I - oligo(dT) primer [5'  
 TGTTACCAATCTGAAGTGGAGCGCGCTTAATTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT7T3 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaudo.

115 a 151 c 172 g 108 t 1 others

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 1.44

Score: 80.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 50.00%

DB: 9

Indels: 0

Gaps: 0

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x AA443351 (1-547)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
 |||||  
 Db 174 GACAAGGGCTGCCCGCGAGAGAGAGCCACCCCTCTGACGTCC 218

## RESULT 15

BM720098

LOCUS

DEFINITION UI-E-E00-ahz-c-05-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone  
 UI-E-E00-ahz-c-05-0-UI 5', mRNA sequence.

ACCESSION BM720098

VERSION BM720098.1 GI:19039119

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 561)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalizaton and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT

Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

## FEATURES

source

1..561

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="UI-E-E00-ahz-c-05-0-UI"

/tissue\_type="fetal eye"

/dev\_stage="fetal"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone\_lib="UI-E-E00"

/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
 modified polylinker; Site1: EcoR I; Site2: Not I;  
 UI-E-E00 is a cDNA library containing the following  
 tissue(s): fetal eye. The library was constructed  
 according to Bonaldo, Lennon and Soares, Genome Research,  
 6:791-806, 1996. First strand cDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded cDNA was ligated to an EcoR I adaptor, digested  
 with Not I, and cloned directionally into pT7T3-Pac  
 vector. The oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tag for this library is CCGTATACC. This library  
 was created for the program, Gene Discovery in the Visual  
 System, supported by National Eye Institute (NEI)."

97 a 171 c 179 g 113 t

1 others

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 1.49

Score: 80.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 50.00%

DB: 12

Indels: 0

Gaps: 0

Length: 561

Matches: 15

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0



JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x BM720098 (1-561)

Qy	1	AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer	15
Db	400	GACAAGGGCTGCCCGCCGAGCAGAGAGAGCCGCCCTCTGACGTCC	444

Search completed: October 16, 2003, 17:03:59  
Job time : 240.794 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 16, 2003, 11:08:07 : Search time 31.0111 Seconds  
(without alignments)  
2540.503 Million cell updates/sec

Title: JUNC\_SEQ3\_SEQ2RES991\_  
Perfect score: 160  
Sequence: 1 DKGPAPQASPLTSQNEGLGPPSPMDSTF 30

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1750203.seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-O/cgn2\_1/USPTO\_spool/p/HOLLERAN480/runat\_15102003\_131915\_20662/app\_query.fasta\_1.4685  
-DB=PublishedApplications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-NAXLEN=2000000000 -USER=HOLLERAN480 -CGN\_1\_1\_397@runat\_15102003\_131915\_20662  
-NCPU=6 -ICPU=3 -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -YGAPOP=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	80	50.0	201	Sequence 3, Appli

2	80	50.0	3765	12	US-10-207-498-5	Sequence 5, Appli
3	80	50.0	3768	9	US-09-811-123-8	Sequence 8, Appli
4	80	50.0	3768	9	US-09-811-115-2	Sequence 2, Appli
5	80	50.0	3768	10	US-09-854-356-9	Sequence 9, Appli
6	80	50.0	3768	10	US-09-930-125-1	Sequence 1, Appli
7	80	50.0	3768	12	US-10-313-644-1	Sequence 1, Appli
8	80	50.0	3771	10	US-09-854-356-11	Sequence 11, Appli
9	80	50.0	3955	10	US-09-870-759-117	Sequence 117, Appli
10	80	50.0	3955	10	US-09-854-356-10	Sequence 10, Appli
11	80	50.0	3955	12	US-09-751-708A-117	Sequence 117, Appli
12	80	50.0	4473	11	US-09-441-411-5	Sequence 5, Appli
13	80	50.0	4473	12	US-10-101-510-81	Sequence 81, Appli
14	80	50.0	4473	14	US-10-146-473-32	Sequence 32, Appli
15	80	50.0	4473	14	US-10-207-655-44	Sequence 44, Appli
16	80	50.0	4530	10	US-08-877-177-11	Sequence 11, Appli
17	80	50.0	4530	12	US-10-007-926A-119	Sequence 119, Appli
18	80	50.0	4530	12	US-10-101-510-124	Sequence 124, Appli
19	80	50.0	4530	12	US-10-338-730-1	Sequence 1, Appli
20	80	50.0	4530	14	US-10-177-293-125	Sequence 125, Appli
21	80	50.0	4543	10	US-09-769-508-1	Sequence 1, Appli
22	80	50.0	4606	12	US-09-971-392-70	Sequence 70, Appli
23	80	50.0	4642	14	US-10-198-846-10896	Sequence 10896, A
24	80	50.0	9274	9	US-09-811-123-7	Sequence 7, Appli
25	80	50.0	9274	9	US-09-811-115-1	Sequence 1, Appli
26	74	46.2	1115	14	US-10-102-806-165	Sequence 165, Appli
27	74	46.2	1713	12	US-10-378-393-14	Sequence 14, Appli
28	74	46.2	1755	10	US-09-930-125-6	Sequence 6, Appli
29	74	46.2	1767	10	US-09-930-125-4	Sequence 4, Appli
30	74	46.2	1773	10	US-09-930-125-7	Sequence 7, Appli
31	74	46.2	1806	10	US-09-930-125-5	Sequence 5, Appli
32	74	46.2	2411	12	US-10-378-393-10	Sequence 10, Appli
33	65	40.6	23432	9	US-09-764-869-1332	Sequence 1332, Ap
34	65	40.6	23432	14	US-10-091-504-1332	Sequence 1332, Ap
35	60	37.5	111282	13	US-10-094-989-3	Sequence 3, Appli
36	59	36.9	4092	11	US-09-925-388-5	Sequence 5, Appli
37	58	36.2	4092	12	US-10-431-846-5	Sequence 5, Appli
38	58	36.2	1938	12	US-10-422-264-25	Sequence 25, Appli
39	58	36.2	2316	12	US-10-422-264-27	Sequence 27, Appli
40	58	36.2	2604	12	US-10-432-264-23	Sequence 23, Appli
41	58	36.2	3320	10	US-09-862-658-1	Sequence 1, Appli
42	58	36.2	3320	14	US-10-175-696-22	Sequence 22, Appli
43	58	36.2	3384	12	US-10-422-264-39	Sequence 29, Appli
44	57.5	35.9	363	11	US-09-918-995-28706	Sequence 28706, A
45	57.5	35.9	2691	10	US-09-967-768A-292	Sequence 292, App

ALIGNMENTS

RESULT 1  
US-10-109-213-3  
; Sequence 3, Application US/10109213  
; Publication No. US20020168670A1  
; GENERAL INFORMATION:  
; APPLICANT: Ecker, David J.  
; TITLE OF INVENTION: Identification of Disease Predictive Nucleic Acids  
; FILE REFERENCE: IBIS0009  
; CURRENT APPLICATION NUMBER: US/10/109,213  
; PRIOR FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: US/09/200,355  
; PRIOR FILING DATE: 1998-11-25  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 201  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-10-109-213-3

Alignment Scores:	0.000792	Length:	201
Pred. No.:	80.00	Matches:	15
Score:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

Query Match: 50.00% Indels: 0  
DB: 13 Gaps: 0

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x US-10-109-213-3 (1-201)

QY 1 AsPLySGlYcysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 116 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCUCUGAGGUCC 160

## RESULT 2

US-10-207-498-5  
; Sequence 5, Application US/10207498  
; Publication No. US20030143568A1  
; GENERAL INFORMATION:  
; APPLICANT: Elizabeth Singer  
; APPLICANT: Ralf Landgraf  
; APPLICANT: Dennis J. Slamon  
; APPLICANT: David Eisenberg  
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING  
; FILE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HEREGULIN AND HER3  
; FILE REFERENCE: 30448.103-US-U1  
; CURRENT APPLICATION NUMBER: US/10/207,498  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 60/308,431  
; PRIOR FILING DATE: 2001-07-27  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 3765  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(3765)  
US-10-207-498-5

Alignment Scores:  
Pred. No.: 0.0194 Length: 3765  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.00% Indels: 0  
DB: 12 Gaps: 0

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x US-10-207-498-5 (1-3765)

QY 1 AsPLySGlYcysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTCC 1959

## RESULT 3

US-09-811-123-8  
; Sequence 8, Application US/09811123  
; Patent No. US20020001587A1  
; GENERAL INFORMATION:  
; APPLICANT: Sharon Erickson  
; APPLICANT: Ralph Schwall  
; APPLICANT: Mark Sliwowski  
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB  
; FILE OF INVENTION: ANTIBODY-MAYTANSINOID CONJUGATES  
; FILE REFERENCE: GENENT.073A2  
; CURRENT APPLICATION NUMBER: US/09/811,123  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/238,327  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 09/602,530  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 3768  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-811-123-8

Alignment Scores:  
Pred. No.: 0.0194 Length: 3768  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.00% Indels: 0  
DB: 9 Gaps: 0

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x US-09-811-123-8 (1-3768)

QY 1 AsPLySGlYcysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTCC 1959

## RESULT 4

US-09-811-115-2  
; Sequence 2, Application US/09811115  
; Patent No. US20020035736A1  
; GENERAL INFORMATION:  
; APPLICANT: Erickson, Sharon  
; APPLICANT: Schwall, Ralph  
; APPLICANT: King, Kathleen  
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL  
; FILE REFERENCE: GENENT.034A  
; CURRENT APPLICATION NUMBER: US/09/811,115  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/189,844  
; PRIOR FILING DATE: 2000-03-16  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO.2  
; LENGTH: 3768  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-811-115-2

Alignment Scores:  
Pred. No.: 0.0194 Length: 3768  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.00% Indels: 0  
DB: 9 Gaps: 0

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x US-09-811-115-2 (1-3768)

QY 1 AsPLySGlYcysProAlaGluGlnArgAlaSerProLeuThrSer 15  
|||||  
Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTCC 1959

## RESULT 5

US-09-854-356-9  
; Sequence 9, Application US/09854356  
; Patent No. US20020177567A1  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Gheysen, Dirk  
; APPLICANT: Corixa Corporation  
; APPLICANT: SmithKline Beecham Biologicals S. A.  
; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
; FILE REFERENCE: 014058-009810PC  
; CURRENT APPLICATION NUMBER: US/09/854,356  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: US 09/493,480  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/117,976  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 3768

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

```

```

Alignment Scores:
Pred. No.: 0.0194 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
Gaps: 0
DB:

```

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x US-09-854-356-9 (1-3768)

```

Qy 1 AsplysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|
Db 1915 GACAAGGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGAGCTCC 1959

```

RESULT 6

```

US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930.125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3765)
US-09-930-125-1

```

```

Alignment Scores:
Pred. No.: 0.0194 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
Gaps: 0
DB:

```

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x US-09-930-125-1 (1-3768)

```

Qy 1 AsplysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|
Db 1915 GACAAGGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGAGCTCC 1959

```

RESULT 7

```

US-10-313-644-1
; Sequence 1, Application US/10313644
; Publication No. US20030157119A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Cheever, Martin A.
; APPLICANT: Hand-Zimmerman, Susan
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
; FILE REFERENCE: 210121.483C3
; CURRENT APPLICATION NUMBER: US/10/313.644
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3765)
US-10-313-644-1

```

```

Alignment Scores:
Pred. No.: 0.0194 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
Gaps: 0
DB:

```

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x US-10-313-644-1 (1-3768)

```

Qy 1 AsplysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|
Db 1915 GACAAGGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGAGCTCC 1959

```

RESULT 8

```

US-09-854-356-11
; Sequence 11, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854.356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 3771
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: mouse HER-2/neu
US-09-854-356-11

```

```

Alignment Scores:
Pred. No.: 0.0195 Length: 3771
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
DB:

```

```
;
; NAME/KEY: CDS
; LOCATION: (26)..(3799)
; OTHER INFORMATION: rat HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (26)..(1990)
; OTHER INFORMATION: extracellular domain (ECD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2057)..(3796)
; OTHER INFORMATION: intracellular domain (ICD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2188)..(3022)
; OTHER INFORMATION: kinase domain (KD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2999)..(3796)
; OTHER INFORMATION: phosphorylation domain (PD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2999)..(3173)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of rat HER-2/neu
US-09-854-356-10

Alignment Scores:
Pred. No.: 0.0205 Length: 3955
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-854-356-11 (1-3771)

QY 16 GlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
|||||
Db 2971 CAGACGAGGAGCTTAGGCCCTCCAGCCCATGGACGACGACCTTC 3015

RESULT 9
US-09-870-759-117
; Sequence 117, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 117
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17)..(3799)
; OTHER INFORMATION:
US-09-870-759-117

Alignment Scores:
Pred. No.: 0.0205 Length: 3955
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-870-759-117 (1-3955)

QY 16 GlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
|||||
Db 2999 CAGACGAGGAGCTTGGGCCCATCCAGCCCATGGACGACGACCTTC 3043

RESULT 10
US-09-854-356-10
; Sequence 10, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: rat HER-2/neu cDNA
```

```
;
; NAME/KEY: CDS
; LOCATION: (26)..(3799)
; OTHER INFORMATION: rat HER-2/neu protein
; NAME/KEY: misc_feature
; LOCATION: (26)..(1990)
; OTHER INFORMATION: extracellular domain (ECD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2057)..(3796)
; OTHER INFORMATION: intracellular domain (ICD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2188)..(3022)
; OTHER INFORMATION: kinase domain (KD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2999)..(3796)
; OTHER INFORMATION: phosphorylation domain (PD) of rat HER-2/neu
; NAME/KEY: misc_feature
; LOCATION: (2999)..(3173)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of rat HER-2/neu
US-09-854-356-10

Alignment Scores:
Pred. No.: 0.0205 Length: 3955
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-854-356-10 (1-3955)

QY 16 GlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
|||||
Db 2999 CAGACGAGGAGCTTGGGCCCATCCAGCCCATGGACGACGACCTTC 3043

RESULT 11
US-09-751-708A-117
; Sequence 117, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 117
; LENGTH: 3955
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17)..(3799)
; OTHER INFORMATION:
US-09-751-708A-117

Alignment Scores:
Pred. No.: 0.0205 Length: 3955
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 12 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-751-708A-117 (1-3955)

QY 16 GlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
|||||
Db 2999 CAGACGAGGAGCTTGGGCCCATCCAGCCCATGGACGACGACCTTC 3043
```



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:06:28 ; Search time 28.4941 Seconds  
(without alignments)  
2842.104 Million cell updates/sec

Title: JUNC\_SEQ3\_SEQ2RES991\_  
Perfect score: 160  
Sequence: 1 DKGCPAEQASPLTSQNEDLCPSPMDSTF 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/Cyn2.1/USPFO\_spool\_p/HOLLERAN480/runat\_15102003\_131912\_20526/app\_query.fasta\_1.4685  
-DB=N.Geneseq\_19Jun03 -QMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTEXT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=HOLLERAN480 @CGN\_1\_1\_1758 @runat\_15102003\_131912\_20526 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	150	93.8	2763	24	ABA92252	Mouse Her-2/neu ex
2	150	93.8	2781	24	ABA92253	Her-2/neu extracel
3	82	51.2	2385	18	AAAT72725	Her2-GM-CSF immuno
4	80	50.0	69	21	AAA14815	DNA encoding a wil
5	80	50.0	201	21	AAAG2125	Her2 normal form t
6	80	50.0	2125	19	AAV21727	Humanised vector p
7	80	50.0	2871	21	AAZ50586	DC8scFv-erbB2EC fu
8	80	50.0	3600	21	AAAS9736	cDNA encoding huma
9	80	50.0	3678	24	ABK86207	HER-2/neu oncogene
10	80	50.0	3768	17	AAAT07339	Human HER-2/neu on
11	80	50.0	3768	20	AAAT01912	Human heregulin 2
12	80	50.0	3768	21	AAA09455	Human HER-2/neu pr
13	80	50.0	3768	22	AAH23392	Human ERBB2 polyinu
14	80	50.0	3768	24	ABZ35744	Human ERBB2 DNA fr
15	80	50.0	3768	24	ABX09987	Human ERBB2 cDNA
16	80	50.0	3768	24	AAAD3935	Human ERBB2 DNA SE
17	80	50.0	3768	24	ABV78168	Human Her-2/neu pr
18	80	50.0	3768	24	AAAD32743	Human Her-2/neu CD
19	80	50.0	3768	24	ABA92250	Human Her-2/neu DN
20	80	50.0	3768	24	ABK10730	Human polynucleoti
21	80	50.0	3768	24	ABL91709	Human HER2 (ERBB2)
22	80	50.0	3768	24	ABK14058	Mouse Her-2/neu CD
23	80	50.0	3771	21	AAA89737	Nucleotide sequenc
24	80	50.0	3771	22	AAH42210	Mouse Her-2/neu CD
25	80	50.0	3771	24	ABA92251	Rat neu promoter.
26	80	50.0	3955	16	AAAT01590	Rat HER-2/neu prot
27	80	50.0	3955	21	AAA97553	Sequence encoding
28	80	50.0	4299	14	AAQ46083	cDNA encoding the
29	80	50.0	4299	14	AAA14812	Human tumour anti
30	80	50.0	4472	21	AAAL4812	HER-2 nucleic acid
31	80	50.0	4473	19	AB076220	Human gene express
32	80	50.0	4473	20	AAZ31071	Human Her-2 DNA
33	80	50.0	4473	24	ABZ34969	Her-2/neu (ERBB2)/c
34	80	50.0	4473	16	AAAT01585	Human HER2 gene.
35	80	50.0	4530	18	AAAT71253	Nucleotide sequenc
36	80	50.0	4530	21	AAZ60815	Human tyrosine kin
37	80	50.0	4530	22	AAJ19731	Human gene express
38	80	50.0	4530	24	ABZ35012	Breast carcinoma r
39	80	50.0	4530	24	ABV94128	Human HER2-neu SEQ
40	80	50.0	4530	24	ABN85585	Human cDNA differe
41	80	50.0	4530	24	ABK83918	Breast cancer asso
42	80	50.0	4530	25	ACC50139	Human Her2/Neu enc
43	80	50.0	4530	25	ABQ83856	HER2 transgene pla
44	80	50.0	9274	22	AAF24237	
45	80	50.0				

ALIGNMENTS

RESULT 1  
ABA92252  
ID ABA92252 standard; cDNA; 2763 BP.  
AC  
AC ABA92252;  
XX  
XX 17-JUN-2002 (first entry)  
DT  
XX Mouse Her-2/neu extracellular-phosphorylation domain fusion cDNA.  
DE  
XX Her-2/neu; oncogene; cancer; tumour; vaccine; tyrosine kinase;  
KW receptor; mouse; gene therapy; gene; ss.  
KW  
XX Mus musculus.  
OS  
XX Key  
FH Location/Qualifiers  
FT CDS 1..2763



QY 1 AsPLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20  
Db 1918 GAACGAGCTGCCACGACAGACAGAGACCCAGTGCAGTCTCAGAACGAGGACTTA 1977  
QY 21 GlyProSerSerProMetAspSerThrPhe 30  
Db 1978 GGCCCTCCAGCCCATGGACAGCACCTTC 2007

## RESULT 3

AAAT72725  
ID AAAT72725 standard; cDNA; 2385 BP.  
XX  
AC AAAT72725;  
XX  
DT 17-SEP-1997 (first entry)  
XX  
DE Her2-GM-CSF immunostimulant fusion protein DNA.  
XX  
KW Her2-GM-CSF; granulocyte macrophage colony stimulating factor;  
KW growth factor receptor; oncogene; immunostimulant; cancer;  
KW therapy; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 11..2359  
FT /tag= a  
FT /product= GM-CSF-Her2 fusion protein  
FT mRNA 11..1969  
FT /tag= b  
FT /product= Her2  
FT mRNA 1970..1975  
FT /tag= c  
FT /product= Leu-Glu linker  
FT mRNA 1976..2359  
FT /tag= d  
FT /product= GM-CSF  
XX  
PN WO9724438-A1.  
XX  
PD 10-JUL-1997.  
XX  
PF 23-DEC-1996; 96WO-US20241.  
XX  
PR 28-DEC-1995; 95US-0579823.  
XX  
PA (ACT1-) ACTIVATED CELL THERAPY INC.  
XX  
PI Laus R, Ruegg CL, Wu H;  
XX  
DR WPI; 1997-363674/33.  
DR P-PSDB; AAW19764.  
XX  
PT Potent APC that activates T-cells to give multivalent cellular  
PT immune response - can also induce a cytotoxic T-cell response in a  
PT vertebrate subject  
XX  
PS Disclosure; Fig 8; 45pp; English.  
XX  
CC A nucleic acid molecule (AAAT72725) codes for a fusion protein  
CC (AAW19764) comprising granulocyte-macrophage colony stimulating  
CC factor (GM-CSF) and Her2, a growth factor receptor that is  
CC over-expressed in breast and ovarian cancer cells. It was  
CC prep'd by PCR amplification of Her2 cDNA from a breast cancer  
CC cell line and fusion to GM-CSF cDNA. Fusion expression vectors can  
CC be used to transfect mammalian and insect cells. The Her2-GM-CSF  
CC fusion protein is used to generate anti-Her2 immunity. Tumour  
CC cells are eliminated by cytotoxic T lymphocytes activated in vivo  
CC or in vitro by exposure to antigen-presenting cells exposed to the  
CC fusion protein.  
XX  
SQ Sequence 2385 BP; 488 A; 780 C; 677 G; 440 T; 0 other;

Alignment Scores:  
Pred. No.: 0.0564 Length: 2385  
Score: 82.00 Matches: 17  
Percent Similarity: 68.00% Conservative: 0  
Best Local Similarity: 68.00% Mismatches: 8  
Query Match: 51.25% Indels: 0  
DB: 18 Gaps: 0

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x AAAT72725 (1-2385)

QY 1 AsPLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20  
Db 1925 GACAAGGCTGCCCGCGCAGACAGAGAGCCAGCCCTCTGACGTCTCTCGAGGCACCCGCC 1984  
QY 21 GlyProSerSerPro 25  
Db 1985 CGCTCGCCAGCCCC 1999

RESULT 4  
AAAT4815  
ID AAAT4815 standard; cDNA; 69 BP.  
XX  
AC AAAT4815;  
XX  
DT 08-AUG-2000 (first entry)  
XX  
DE DNA encoding a wild type erbB-2 receptor protein fragment.  
XX  
KW SPLICE erbB-2 receptor protein; cell transformation disorder; cancer;  
KW tumor cell proliferation; tissue degeneration; arthropathy;  
KW bone resorption; inflammatory disease; degenerative disorder;  
KW wound healing; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200020579-A1.  
XX  
PD 13-APR-2000.  
XX  
PF 01-OCT-1999; 99WO-CA00912.  
XX  
PR 02-OCT-1998; 98US-0165192.  
XX  
PA (UYMC-) UNIV MCMASTER.  
XX  
PI Muller WJ, Siegel PM;  
XX  
DR WPI; 2000-303768/26.  
DR P-PSDB; AAY84681.  
XX  
PT Nucleic acid encoding an erbB 2 receptor protein designated SPLICE  
PT erbB-2, inhibitors of the protein are useful for treatment of cancer -  
XX  
PS Example 2; Fig 3B; 60pp; English.  
XX  
CC The present sequence encodes a peptide of wild type erbB-2 receptor  
CC protein, where the splicing event occurs. An alternatively spliced  
CC variant of erbB-2 receptor protein, designated SPLICE erbB-2 receptor  
CC protein, has an in-frame deletion of 16 amino acids, 2 of which are  
CC conserved cysteine residues, compared to the unspliced protein. The  
CC erbB-2 polynucleotide is used to construct probes for detecting  
CC disorders of cell transformation such as cancer. Antibodies to the  
CC protein may be used to detect SPLICE erbB-2 in a sample. Agents  
CC (e.g. antisense oligonucleotides) which inhibit the expression of  
CC SPLICE erbB-2 are useful for reducing tumor cell proliferation and  
CC treating cancer. Substances which stimulate SPLICE erbB-2 are useful  
CC for treating conditions involving damaged cells including conditions in  
CC which degeneration of tissue occurs, such as arthropathy, bone  
CC resorption, inflammatory diseases, degenerative disorders of the  
CC central nervous system and wound healing.  
XX  
SQ Sequence 69 BP; 13 A; 26 C; 20 G; 10 T; 0 other;



DE Humanised vector pITL-hHER/neu.  
 XX Vector; vaccine; tumour; antigen; plasmid pITL-hHER/neu;  
 XX human; HER-2/neu; C-erbB-2; breast cancer; ds.  
 KW  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Escherichia coli.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 13..921  
 FT /tag= a  
 FT /product= human HER-2/neu  
 FT polyA\_site 922..1181  
 FT /tag= b  
 FT /note= "combined splice and polyA sequences"  
 FT CDS 1195..1401  
 FT /tag= c  
 FT /note= "SupF gene"  
 FT 1412..1864  
 FT /tag= d  
 FT /note= "ColE1 origin of replication"  
 FT promoter 1877..2125  
 FT /tag= e  
 FT /note= "RANTES promoter"  
 FT  
 FT WO9806863-A1.  
 PN  
 XX  
 XX 19-FEB-1998.  
 PD  
 XX 14-AUG-1997; 97WO-US14306.  
 PF  
 XX 14-AUG-1996; 96US-0023931.  
 PR  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX Nelson EL, Nelson PJ;  
 PI  
 XX WPI; 1998-159552/14.  
 DR  
 XX Humanised polynucleotide vectors - comprising human derived promoter  
 PT and sequence acceptance site, used for the production of vaccines  
 XX  
 XX Example 7; Page 41-42; 125pp; English.  
 PS  
 XX Plasmid pITL-hHER2/neu comprises base vector pITL (see AAV21724) and  
 CC a human HER-2/neu nucleic sequence. pITL-hHER2/neu was used to  
 CC evaluate the toxicity of anti-tumour vaccination in rats, and in  
 CC phase I and phase II trials to evaluate polynucleotide vaccination  
 CC in advanced breast cancer. Novel humanised vectors, which can be  
 CC based on pITL, comprise a human-derived promoter or mammalian  
 CC homologue which is functional in mammalian target tissue and cells  
 CC and a sequence acceptance site which accepts cDNA products from  
 CC RT-PCR cloning. The vectors are non-replicating in mammalian cells  
 CC but are capable of extended stable expression of the target  
 CC sequence, generating an immune response in immunised individuals.  
 CC The vectors selectively elicit immune responses to the target  
 CC sequences with little or no immune response to the other components  
 CC of the vectors.  
 XX  
 SQ Sequence 2125 BP; 449 A; 650 C; 598 G; 428 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.0956 Length: 2125  
 Score: 80.00 Matches: 15  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.00% Indels: 0  
 DB: 19 Gaps: 0  
 JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x AAV21727 (1-2125)  
 Oy 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

Db 112 GACAAGGGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCC 156  
 RESULT 7  
 AAZ50586  
 ID AAZ50586 standard; DNA; 2871 BP.  
 XX  
 AC AAZ50586;  
 XX  
 DT 23-MAY-2000 (first entry)  
 XX  
 DE DC8scFv-erbB2EC fusion construct containing tetramerisation domain.  
 DE  
 KW DC8 scFv; single-chain variable fragment; erbB2EC; extracellular domain;  
 KW human; fusion construct; tetramerisation domain; constant domain;  
 KW heteromilbody; multifunctional compound; melanoma; sarcoma;  
 KW immunoglobulin; cytostatic; immunostimulatory; antileukaemia; diagnosis;  
 KW antiproliferative; prevention; treatment; malignant; haematopoietic cell;  
 KW lymphoma; leukaemia; solid tumour; carcinoma; ds.  
 KW  
 OS Chimeric - Unidentified.  
 OS Chimeric - Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 10..2865  
 FT /tag= a  
 FT /product= "DC8scFv-erbB2EC fusion protein"  
 FT sig\_peptide 10..66  
 FT /tag= b  
 FT /label= Leader\_sequence  
 FT misc\_feature 67..390  
 FT /tag= c  
 FT /label= DC8scFv\_light\_chain\_variable\_region  
 FT misc\_feature 391..435  
 FT /tag= d  
 FT /label= Glycine-Serine-linker\_DNA  
 FT misc\_feature 436..771  
 FT /tag= e  
 FT /label= DC8scFv\_heavy\_chain\_variable\_region  
 FT misc\_feature 775..807  
 FT /tag= f  
 FT /note= "5' end of human IgG3 upper hinge region  
 FT with 3 additional nucleotides"  
 FT misc\_feature 808..924  
 FT /tag= g  
 FT /label= Human\_p53\_tetramerisation\_domain  
 FT misc\_feature 925..945  
 FT /tag= h  
 FT /label= Short\_peptide\_linker  
 FT misc\_feature 946..2844  
 FT /tag= i  
 FT /label= erbB2EC\_domain  
 FT misc\_feature 2845..2862  
 FT /tag= j  
 FT /label= His\_tag  
 XX WO200006605-A2.  
 XX  
 PD 10-FEB-2000.  
 XX  
 XX 28-JUL-1999; 99WO-EP05416.  
 XX  
 XX 28-JUL-1998; 98EP-0114082.  
 PR  
 XX (MICR-) MICROMET GES BIOMEDIZINISCHE FORSCHUNG.  
 PA  
 XX Kufer P, Dreier T, Baeuerle PA, Borschert K, Zettl F;  
 PI  
 XX WPI; 2000-195265/17.  
 XX P-PSDB; AAY44993.  
 DR  
 XX New multifunctional compounds useful for preventing and/or treating  
 PT malignant cell growth and for detection and diagnosis -





PN US5869445-A.  
XX  
PD 09-FEB-1999.  
XX  
XX 01-APR-1996; 96US-0625101.  
XX  
XX 01-APR-1996; 96US-0625101.  
PR 17-MAR-1993; 93US-0033644.  
PR 12-AUG-1993; 93US-0106112.  
PR 31-MAR-1995; 95US-0414417.  
XX  
XX (UNIW ) UNIV WASHINGTON.  
XX  
XX Cheever MA, Disis ML;  
PI WPI; 1999-152835/13.  
DR P-PSDB; AAW92406.  
XX  
XX Use of HER-2/neu polypeptides - for eliciting an immune response to  
PT an HER-2/neu associated malignancy, particularly for treating or  
PT preventing tumours  
XX  
XX Claim 1a; Column 23-32; 26pp; English.  
XX  
XX This sequence encodes the human HER-2/neu oncogene protein. A fragment  
CC of this protein is used in a method for eliciting or enhancing an immune  
CC response to HER-2/neu protein. The polypeptide can stimulate T cells and  
CC B cells to produce an immune response to the HER-2/neu protein. The  
CC method can be used for immunisation against a malignancy in which the  
CC HER-2/neu oncogene is associated and in the treatment of an existing  
CC tumour, or to prevent tumour occurrence or reoccurrence.  
XX  
SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;  
  
Alignment Scores: Pred. No.: 0.192 Length: 3768  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.00% Indels: 0  
DB: 20 Gaps: 0  
  
JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x AAX01912 (1-3768)  
QY 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15  
|||||  
Db 1915 GACAGGGCTGCCCGCCGAGCAGAGCCAGCCCTCTGACGTC 1959  
  
RESULT 12  
AAA09455  
ID AAA09455 standard; DNA; 3768 BP.  
XX  
AC AAA09455;  
XX  
XX 10-AUG-2000 (first entry)  
DE Human heregulin 2 (Her2) coding sequence.  
XX  
XX Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;  
KW self-protein; cell-associated peptide antigen; foreign epitope;  
KW cancer; breast cancer; prostate cancer; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200020027-A2.  
PN  
XX 13-APR-2000.  
PD  
XX 05-OCT-1999; 99WO-DK00525.  
XX  
XX 05-OCT-1998; 98DK-0001261.  
PR  
XX 20-OCT-1998; 98US-0105011.  
XX

PA (MEBI-) M & E BIOTECH AS.  
XX  
XX Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;  
PI Gautam A, Birk P, Karlsson G;  
XX  
XX WPI; 2000-349917/30.  
DR P-PSDB; AAY92620.  
XX  
XX Inducing immune responses to weakly immunogenic, tumor associated  
PT peptide antigens for the treatment of breast and prostate cancer  
XX  
XX Claim 62; Page 187-193; 220pp; English.  
XX  
XX The claims detail a method for inducing immune responses against weakly  
CC immunogenic cell-associated peptide antigens (PA) such as those  
CC associated with cancers (i.e. self-proteins), for example, human  
CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or  
CC fibroblast growth factor 8b (FGF8b). The method comprises effecting  
CC simultaneous presentation by antigen producing cells (APCs) of the  
CC animals immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)  
CC group derived from the PA and/or at least 1 B-cell group derived from the  
CC cell-associated PA; and (2) at least 1 first T helper cell group which is  
CC foreign to the animal. Analogues of human PSM, human Her2 and  
CC human/murine FGF8b comprising a substantial part of all known and  
CC predicted CTL and B-cell epitopes of the respective PA and including at  
CC least one foreign T helper epitope are also claimed. The method is used  
CC to treat prostate, prostate/breast or breast cancer when the PA is human  
XX PSM, FGF8b and Her2, respectively.  
XX  
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;  
  
Alignment Scores: Pred. No.: 0.192 Length: 3768  
Score: 80.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.00% Indels: 0  
DB: 21 Gaps: 0  
  
JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x AAA09455 (1-3768)  
QY 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15  
|||||  
Db 1915 GACAGGGCTGCCCGCCGAGCAGAGCCAGCCCTCTGACGTC 1959  
  
RESULT 13  
AAH23392  
ID AAH23392 standard; DNA; 3768 BP.  
XX  
AC AAH23392;  
XX  
XX 25-SEP-2001 (first entry)  
DT  
XX Human HER-2/neu protein encoding DNA.  
DE  
XX Antigen-presenting cell; immunogenic; immune response; HER-2/neu;  
KW oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.  
XX  
XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
FH 1..3768  
FT CDS /\*tag= a  
FT /\*product= "HER-2/neu protein"  
XX  
XX WO200153463-A2.  
PN  
XX 26-JUL-2001.  
PD  
XX 19-JAN-2001; 2001WO-US01850.  
PF  
XX 21-JAN-2000; 2000US-0177545.  
XX





CC (ds) structure of not more than 49 consecutive nucleotides (nt), where  
CC at least a segment of one strand of the ds structure is complementary  
CC with the target gene and the cells are treated with interferon before  
CC introduction of dsRNA1. The method is used to inhibit expression of  
CC target genes, particularly oncogenes, cytokine genes, Id (not defined)  
CC protein genes; developmental or prion genes, or genes expressed in  
CC pathogenic organisms (particularly plasmodia) or in viruses or viroids  
CC (pathogenic in humans, animals or plants). Treating the cells with  
CC interferon greatly increases the extent to which dsRNA can inhibit  
CC expression of the target genes, and the effect is even greater when dsRNA  
CC are modified to increase their stability. ABX0936-ABX10075 represent  
CC gene fragments used to illustrate the method of the invention.

XX

SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:

Pred. No.:	0.192	Length:	3768
Score:	80.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	50.00%	Indels:	0
DB:	24	Gaps:	0

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x ABX09987 (1-3768)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15

Db 1915 GACAAGGGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTCC 1959

Search completed: October 15, 2003, 23:48:35  
Job time : 33.4941 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 15, 2003, 21:09:13 ; Search time 432.026 Seconds  
(without alignments)  
2840.777 Million cell updates/sec

Title: JUNC\_SEQ3\_SEQ2RES991\_  
Perfect score: 160  
Sequence: 1 DKGCPAEORASPLTSQNEIDLGSPSPMDSTF 30

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n\_model -DEV=xlp  
-Q/csn2\_1/USPTO\_pool\_P/HOLLERAN480/runat\_15102003\_131912\_20536/app\_query.fasta\_1.4685  
-DB=GenEmbl -QFMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=HOLLERAN480 -SCGN\_1\_1\_22724\_@runat\_15102003\_131912\_20536 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_hgt: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vi: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_md: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*

29: em\_vi: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_inv: \*  
32: em\_htg\_other: \*  
33: em\_htg\_mus: \*  
34: em\_htg\_pln: \*  
35: em\_htg\_rod: \*  
36: em\_htg\_mam: \*  
37: em\_htg\_vrt: \*  
38: em\_sy: \*  
39: em\_htgo\_hum: \*  
40: em\_htgo\_mus: \*  
41: em\_htgo\_other: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	150	93.8	2763	6	AX380942	AX380942 Sequence
2	150	93.8	2781	6	AX380944	AX380944 Sequence
C 3	91.5	57.2	186134	10	AL591390	AL591390 Mouse DNA
C 4	91.5	57.2	214019	2	AC064803	AC064803 Mus muscu
5	82	51.2	2385	6	AR082744	AR082744 Sequence
6	82	51.2	2385	6	AR099963	AR099963 Sequence
7	82	51.2	2385	6	AR143949	AR143949 Sequence
8	80	50.0	153	6	AX022984	AX022984 Sequence
9	80	50.0	201	6	AR229723	AR229723 Sequence
10	80	50.0	2083	10	BC027080	BC027080 Mus muscu
11	80	50.0	2871	6	AX023363	AX023363 Sequence
12	80	50.0	3678	6	AX050514	AX050514 Sequence
13	80	50.0	3768	6	AR034479	AR034479 Sequence
14	80	50.0	3768	6	AX060704	AX060704 Sequence
15	80	50.0	3768	6	AX201817	AX201817 Sequence
16	80	50.0	3768	6	AX380923	AX380923 Sequence
17	80	50.0	3768	6	AX384604	AX384604 Sequence
18	80	50.0	3768	6	AX465456	AX465456 Sequence
19	80	50.0	3768	6	AX467229	AX467229 Sequence
20	80	50.0	3768	6	AX481438	AX481438 Sequence
21	80	50.0	3771	6	AX189662	AX189662 Sequence
22	80	50.0	3771	6	AX380925	AX380925 Sequence
23	80	50.0	3955	6	AX380924	AX380924 Sequence
24	80	50.0	3955	6	I21129	I21129 Sequence 14
25	80	50.0	3955	6	I59750	I59750 Sequence 14
26	80	50.0	3955	10	RNNEUR	X03362 Rat mRNA fo
27	80	50.0	4473	6	AR080259	AR080259 Sequence
28	80	50.0	4473	6	AR167390	AR167390 Sequence
29	80	50.0	4473	9	HSEB2R	X03363 Human c-erb
30	80	50.0	4530	6	AR202597	AR202597 Sequence
31	80	50.0	4530	6	AR283481	AR283481 Sequence
32	80	50.0	4530	6	AX282577	AX282577 Sequence
33	80	50.0	4530	6	AX587649	AX587649 Sequence
34	80	50.0	4530	6	AX644071	AX644071 Sequence
35	80	50.0	4530	6	BD005474	BD005474 Cellular
36	80	50.0	4530	6	I21124	I21124 Sequence 9
37	80	50.0	4530	6	I59745	I59745 Sequence 9
38	80	50.0	4530	9	HUMHER2A	M11730 Human tyros
39	80	50.0	4694	10	BC046811	BC046811 Mus muscu
40	80	50.0	4695	10	BC053078	BC053078 Mus muscu
41	80	50.0	4727	10	AY116182	AY116182 Rattus no
42	80	50.0	9274	6	AX060703	AX060703 Sequence
43	80	50.0	13450	9	AB096612	AB096612 Homo sapi
44	80	50.0	16572	9	AB096614	AB096614 Homo sapi
45	80	50.0	20271	9	AB096613	AB096613 Homo sapi

# ALIGNMENTS

RESULT 1

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x AX380944 (1-2781)

Qy 1 AsplyGlyCysProAlaGluGlnAInrGAtAserProLeuThrSerGlnAsnGluAspLeu 20  
:::.....|:::.....|:::.....|:::.....|:::.....|:::.....|:::.....|:::.....|

Db 1918 GAACGAGGCTGCCACGACGACGACGACGACGACGACGACGACGACGCTCTCAGAACGAGACTTA 1977

Qy 21 GlyProSerSerProMetAspSerThrPhe 30  
|||.....|.....|.....|.....|.....|.....|.....|.....|

Db 1978 GGCCCCCTCCAGCCCCCATGGACAGCACCTTC 2007

RESULT 3

AL591390/c  
LOCUS  
DEFINITION  
complete sequence.

ACCESSION AL591390  
VERSION AL591390.8 GI:17017767  
KEYWORDS HTG.  
SOURCE Mus musculus  
ORGANISM Mus musculus (house mouse)  
  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 186134)  
Phillimore,B.  
Direct Submission  
Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Nov 20, 2001 this sequence version replaced gi:16535512.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em: EMBL; Sw:  
SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP  
database can be found at  
http://www.sanger.ac.uk/projects/C\_elegans/wormpep RP23-355L10 is  
from the RPCI-23 Mouse PAC Library  
constructed by the group of Pieter de Jong.  
For further details see http://www.chori.org/bacpac/home.htm  
VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone  
RP23-355L10 It may be shorter because we sequence overlapping  
sections only once, except for a short overlap.  
The true right end of clone RP23-355L10 is at 136134 in this  
sequence. The true right end of clone RP23-438D7 is at 2000 in this  
sequence.

FEATURES  
source Location/Qualifiers  
1..186134  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosomes="11"  
/clone="RP23-355L10"  
/clone\_lib="RPCI-23"  
17991..18044  
/note="Sequence from uni-directional primer reads and dGTP  
big dye terminator reads only."

misc\_feature  
47794..47812  
/note="Sequence from uni-directional primer reads and dGTP  
big dye terminator reads only."

misc\_feature  
46287 a 47900 c 47689 g 44258 t

BASE COUNT  
ORIGIN



misc\_feature 10885..122875  
 /note="assembly\_fragment"  
 misc\_feature 122976..139026  
 /note="assembly\_fragment"  
 misc\_feature 139127..159729  
 /note="assembly\_fragment"  
 misc\_feature 159830..178173  
 /note="assembly\_fragment"  
 misc\_feature 178274..200970  
 /note="assembly\_fragment"  
 misc\_feature 201071..214019  
 /note="assembly\_fragment"  
 clone\_end:T7  
 vector\_side:right"  
 BASE COUNT 52809 a 53348 c 53260 g 53244 t 1358 others.  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.758 Length: 214019  
 Score: 91.50 Matches: 20  
 Percent Similarity: 72.41% Conservative: 1  
 Best Local Similarity: 68.97% Mismatches: 5  
 Query Match: 57.19% Indels: 3  
 DB: 2 Gaps: 1

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x AC064803 (1-214019)

QY 2 LysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGly 21  
 Db 53608 GAGGGCTGT-----CACCTCGTGTCTCCCTTCCTTAGAACGAGGACTTAGGC 53558  
 QY 22 ProSerSerProMetAspSerThrPhe 30  
 Db 53557 CCCTCCAGCCCATGACACCACTTC 53531

RESULT 5  
 AR082744  
 LOCUS AR082744 2385 bp DNA linear PAT 01-SEP-2000  
 DEFINITION Sequence 3 from patent US 5976546.  
 ACCESSION AR082744  
 VERSION AR082744.1 GI:10009534  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 2385)  
 AUTHORS Laus,R., Ruegg,C.Landon. and Wu,H.  
 TITLE Immunostimulatory compositions  
 JOURNAL Patent: US 5976546-A 3 02-NOV-1999;  
 FEATURES Location/Qualifiers  
 source 1..2385  
 /organism="unknown"  
 BASE COUNT 488 a 780 c 677 g 440 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.165 Length: 2385  
 Score: 82.00 Matches: 17  
 Percent Similarity: 68.00% Conservative: 0  
 Best Local Similarity: 68.00% Mismatches: 8  
 Query Match: 51.25% Indels: 0  
 DB: 6 Gaps: 0

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x AR082744 (1-2385)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20  
 Db 1925 GACAAGGGCTGCCCGCCGAGCAGAGCCAGCCCTCTGACGTCCTCGAGGCACCCGCC 1984  
 QY 21 GlyProSerSerPro 25  
 Db 1985 CGCTCGCCGAGCCCC 1999

RESULT 6  
 AR099963  
 LOCUS AR099963 2385 bp DNA linear PAT 14-FEB-2001  
 DEFINITION Sequence 3 from patent US 6080409.  
 ACCESSION AR099963  
 VERSION AR099963.1 GI:12810411  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 2385)  
 AUTHORS Laus,R., Ruegg,C.Landon. and Wu,H.  
 TITLE Immunostimulatory method  
 JOURNAL Patent: US 6080409-A 3 27-JUN-2000;  
 FEATURES Location/Qualifiers  
 source 1..2385  
 /organism="unknown"  
 BASE COUNT 488 a 780 c 677 g 440 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.165 Length: 2385  
 Score: 82.00 Matches: 17  
 Percent Similarity: 68.00% Conservative: 0  
 Best Local Similarity: 68.00% Mismatches: 8  
 Query Match: 51.25% Indels: 0  
 DB: 6 Gaps: 0

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x AR099963 (1-2385)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20  
 Db 1925 GACAAGGGCTGCCCGCCGAGCAGAGCCAGCCCTCTGACGTCCTCGAGGCACCCGCC 1984  
 QY 21 GlyProSerSerPro 25  
 Db 1985 CGCTCGCCGAGCCCC 1999

RESULT 7  
 AR143949  
 LOCUS AR143949 2385 bp DNA linear PAT 08-AUG-2001  
 DEFINITION Sequence 3 from patent US 6210662.  
 ACCESSION AR143949  
 VERSION AR143949.1 GI:15105816  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 2385)  
 AUTHORS Laus,R., Ruegg,C.Landon. and Wu,H.  
 TITLE Immunostimulatory composition  
 JOURNAL Patent: US 6210662-A 3 03-APR-2001;  
 FEATURES Location/Qualifiers  
 source 1..2385  
 /organism="unknown"  
 BASE COUNT 488 a 780 c 677 g 440 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.165 Length: 2385  
 Score: 82.00 Matches: 17  
 Percent Similarity: 68.00% Conservative: 0  
 Best Local Similarity: 68.00% Mismatches: 8  
 Query Match: 51.25% Indels: 0  
 DB: 6 Gaps: 0

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x AR143949 (1-2385)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20  
 Db 1925 GACAAGGGCTGCCCGCCGAGCAGAGCCAGCCCTCTGACGTCCTCGAGGCACCCGCC 1984  
 QY 21 GlyProSerSerPro 25

```

Db      1985 CGTCGCCAGCCGCC 1999
|||||
RESULT 8
LOCUS   AX022984
DEFINITION Sequence 4 from Patent EP0919622.
ACCESSION AX022984
VERSION   AX022984.1 GI:10046479
KEYWORDS
SOURCE   unidentified
ORGANISM unidentified
REFERENCE
1 Springer,C.J. and Marais,R.
TITLE Surface expression of enzyme in gene directed prodrgug therapy
JOURNAL Patent: EP 0919622-A 4 02-JUN-1999;
CANCER RES CAMPAIGN TECH (CB)
FEATURES
source
1. 153
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 27 a 43 c 51 g 32 t
ORIGIN
Alignment Scores:
Pred. No.: 0.0198 Length: 153
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
Gaps: 0
DB: 6

JUNC_SEQ3_SEQ2RES991_ (1-30) x AX022984 (1-153)
Qy 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 10 GACAAGGCTGCCCCCGCGAGCAGAGCCAGCCCTCTGAGCTCC 54

RESULT 9
LOCUS   AR229723
DEFINITION Sequence 3 from patent US 6451524.
ACCESSION AR229723
VERSION   AR229723.1 GI:27269547
KEYWORDS
SOURCE   Unknown.
ORGANISM Unclassified.
REFERENCE
1 (bases 1 to 201)
AUTHORS Eckert,D.J.
TITLE Identification of disease predictive nucleic acids
JOURNAL Patent: US 6451524-A 3 17-SEP-2002;
FEATURES
source
1. 201
/organism="unknown"
BASE COUNT 34 a 67 c 56 g 44 t
ORIGIN
Alignment Scores:
Pred. No.: 0.026 Length: 201
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
Gaps: 0
DB: 6

JUNC_SEQ3_SEQ2RES991_ (1-30) x AR229723 (1-201)
Qy 1 AsplysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 116 GACAAGGCTGCCCCCGCGAGCAGAGCCAGCCCTCTGAGCTCC 160

```

# RESULT 10 LOCUS DEFINITION

BC027080 2083 bp mRNA linear ROD 16-APR-2003  
Mus musculus v-erb-b2 erythroblastic leukemia viral oncogene  
homolog 2, neuro/glioblastoma derived oncogene homolog (avian),  
mRNA (CDNA clone MGC:38648 IMAGE:5356166), complete cds.

BC027080  
BC027080.1 GI:20071980

## ACCESSION VERSION KEYWORDS SOURCE ORGANISM

MGC.  
Mus musculus (house mouse)  
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.  
1 (bases 1 to 2083)

## REFERENCE AUTHORS

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,  
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,  
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,  
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,  
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalius,D.E.,  
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
23388257  
12477932  
2 (bases 1 to 2083)  
Strausberg,R.  
Direct Submission  
Submitted (04-APR-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Jeffrey Green M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lounseged, H.,  
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 54 Row: c Column: 18  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis.

Location/Qualifiers.

1. .2083

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="MGC:38648 IMAGE:5356166"

## FEATURES source

```

/tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI_CGAP_Mam6"
/lab_host="DH10B"
/notes=vector: pCMV-SPORT6"
1..2083
/gene="Erbb2"
/notes="synonyms: Neu, HER-2, c-erbB2, HER2, c-neu"
/db_xref="LocusID:13866"
/db_xref="MGI:95410"
459..1562
/codon_start=1
/product="Erbb2 protein"
/protein_id="AAH27080.1"
/db_xref="GI:20071981"
/db_xref="LocusID:13866"
/translation="MALESILRRRFTHQSDVWSYGVTVWELMTFGAKPYDGIPAREIP
DLLEKERLPQPICITDIYIMVIMVKWMDISECRPRELVSFESRMARDPQRFVVIQ
NEDLGSPSPMDSTFYRSLLDDMDGELVDAEYLVPOQGFSPDPALGTGSTARHRH
SSSARSGGELTGLLEPSEEPSPRLAPSEGAGSDVFDGLAVGTGQLSLSPHDL
SPLOYSDEPTLPLPETDGYVAPLACSPQEVYNQPEVRPOSPLTPEGPPPIRAG
ATLERPKTLPCKNGVKDYFARGGAVENPEYLA PRAGTASQPHSPAPSPAFDNLYY
WQNSSEQPPSPSTFGTETAENPEYLGIDVPV"
BASE COUNT      471 a  588 c  574 g  450 t
ORIGIN
Alignment Scores:
Pred. No.:      0.27      Length:      2083
Score:          80.00      Matches:      15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      50.00%      Indels:      0
DB:              10      Gaps:      0

JUNC_SEQ3_SEQ2RES991_ (1-30) x BC027080 (1-2083)
QY      16 GlnAsnGluAspLeuGlyProSerProMetAspSerThrPhe 30
Db      762 CAGAACGAGGACTTAGGCCCTCCAGCCCATGACAGCACCTC 806

RESULT 11
AX023363
LOCUS      2871 bp      DNA      linear      PAT 15-SEP-2000
DEFINITION      Sequence 34 from Patent WO0006605.
ACCESSION      AX023363
VERSION      AX023363.1      GI:10183775
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Kufner,P., Zettl,F., Dreier,T., Baeuerle,P.A. and Borschert,K.
Heteromabinobodies
Patent: WO 0006605-A 34 10-FEB-2000;
KUFER PETER (DE) ; ZETTL FLORIAN (DE) ; DREIER TORSTEN (DE) ;
BAEUERLE PATRICK A (DE) ; BORSCHERT KATRIN (DE) ; MICROMET GES FUER
BIOMEDIZINIS (DE)
FEATURES
source      Location/Qualifiers
1..2871
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
10..2865
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC08834.1"
/db_xref="GI:10183776"
/translation="MGWSCILFLVATATGVSHSDIOLTSOPATMSAPGKVTMTCTCA
SSVSSTSLHWYQKSGAPKLIYXSTSLASGVPPRSGSGSCTSYSLTISVEARD
AAYIYQQISGYFTTGGGTGLKLEIGGGGGGGGGGGGGGQVRLQSGEIVKPGASV
KLSCASGYLTLSYDHLHWKQWPGRLGWIIGRIDPNSGGTKYDEKFKSKATLTVDKPS
STAYMLQSLTSDSAVYVCARWDYWGQGTTVTVSSGTTPLGDTTHTSGKPLDGEYFTL

```

```

QIRGRERFERELNEALELKDAQAKGEPGGGGGGTQVCTGTMKRLPASPETHLI
DMLRLHYQCOVVQGNLETYLPTNASLFLQDIOBVGQVLIHAHQVQVPLQRLRI
VRGTOLFENYALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLTLTKGGLVLIQRP
QICYQDTILWKDIFHKNNQALTLTIDNRSRACHPCSPCKGSRGCGESSEDQSLTR
TVAGGACRCKGPTDCCHEQCAAGCTGPKHSDCLACLHFNHSGICELRCPALVTYN
TDTFESMPNPEGRYTFEGASCTACPYNTLSDVGSCTLYCELHNSQVTAEDGTQCEK
CSKPCARVCYGLGMEHLREVRVAVTSANIQEFAGCKKIFGSLAFIPESFDGDPASNTAP
LOPEQLOVFTLEETITGYLYISAWPDSLPLDSVFQNLQVIRGRILHNGAYSLTLOGIG
ISWLGRLSRELGSGLALIHNTLHLCFVTPMDOLFNRPHQALLHTANRPEDECYGE
GLACHQLCARGHCWGPGTQCVCNCSQFLRGQECVEBCRVLQGLPRIYVNRHCLPCHP
ECOPQNSVTCFGEADQVACAHYKPPFCVARCPGKSVGLKPLSYMPIWKFPDEEGAC
QPCPINCTHSCVDLDDKQCPAEQASPLTSGHHHHH"
BASE COUNT      598 a  868 c  834 g  571 t
ORIGIN
Alignment Scores:
Pred. No.:      0.372      Length:      2871
Score:          80.00      Matches:      15
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      50.00%      Indels:      0
DB:              6      Gaps:      0

JUNC_SEQ3_SEQ2RES991_ (1-30) x AX023363 (1-2871)
QY      1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db      2797 GACAAGGGCTGCCCGCCGAGCAGACGACGACGACGCTCTGACGTC 2841

RESULT 12
AX0505114
LOCUS      3678 bp      DNA      linear      PAT 27-SEP-2002
DEFINITION      Sequence 7 from Patent WO0240059.
ACCESSION      AX0505114
VERSION      AX0505114.1      GI:23386421
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Mincheff,M.S., Loukinov,D.I. and Zoubak,S.
Methods and compositions for inducing cell-mediated immune
responses
Patent: WO 0240059-A 7 23-MAY-2002;
American Foundation for Biological Research Inc. (US) ; Mincheff,
Milcho S. (US) ; Loukinov, Dmitri I. (US) ; Zoubak, Serguei (US)
FEATURES
source      Location/Qualifiers
1..3678
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
7..3678
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD48804.1"
/db_xref="GI:23386422"
/translation="MRLPASPETHLDMLRLHYQCOVVQGNLETLPTNASLSFLQD
IQEVQVLIHAHQVQVPLQRLRIYRGQLFEDNLTALAVLDNGDPLNNTPTVTGASP
GGLRELQLRSLTEILKGGVLIQRPOLCYODTILWKDIFHKNNQALTLTIDNRSRAC
HPCSPCKGSRGCGESSEDQSLTRTVAGGACRCKGPIPTDCCHEQCAAGCTGPKRS
DCLACLHFNHSGICELRCPALVTYNTPDTPESMPNPEGRYTFEGASCTACPYNTLSDV
GSCTLYCELHNSQVTAEDGTQCEKSKPCARVCYGLGMEHLREVRVAVTSANIQEFAG
CKKIFGSLAFIPESFDGDPASNTAPLQPELVFTELETGYLYISAWPDSLPLDSV
FQNLQVIRGRILHNGAYSLTLOGIGISWLGRLSRELGSGLALIHNTLHLCFVHTVPW
DOLFRNPHQALLHTANRPEDECYGEGLACHQLCARGHCWGPGTQCVCNCSQFLRGQEC
VEBCRVLQGLPRIYVNRHCLPCHPGQNSVTCFGEADQVACAHYKPPFCVARCPGKSV
RCPSGVGPDLSPYMPIWKFPDEEGACQCPINCTHSCVDLDDKQCPAEQASPLTSGV
AVVGIILVYGVVGFGLIKRROQKIRKTYRRELLQVTELVPELTPSGAMPNQAOMRI
LKRETELKRVKVLGSGAGFYVKGWIIPDGENYKIPVAIKVLRNTSPKANKELDEAY
VMAGVSPYVSRLLIGLCTSTVOLTPYGLCDLHVRNRLGSGDQLLNMCMTA
KMGNSYLEDVRLRFTHQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKERLPQPI
MALESILRRRFTHQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKERLPQPI

```





ORIGIN

Alignment Scores:                      Length:                      3768  
 Pred. No.:                      0.488                      Matches:                      15  
 Score:                      80.00                      Conservative:                      0  
 Percent Similarity:                      100.00%                      Mismatches:                      0  
 Best Local Similarity:                      100.00%                      Indels:                      0  
 Query Match:                      50.00%                      Gaps:                      0  
 DB:                      6

JUNC\_SEQ3\_SEQ2RES991\_ (1-30) x AX201817 (1-3768)

QY                      1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15  
 |||||  
 Db                      1915 GACAAGGGCTGCCCCCGCGAGAGAGAGCCCTCTGACGTCC 1959

Search completed: October 16, 2003, 10:58:43  
 Job time : 468.208 secs